THE UNIVERSITY
OF MICHIGAN

SEP 12 1962

ENGINEERING
LIBRARY

Vol. 22, No. 1, January, 1961
Translation Published August, 1961

АВТОМАТИКА
и
ТЕЛЕМЕХАНИКА

SOVIET INSTRUMENTATION AND
CONTROL TRANSLATION SERIES

Automation and Remote Control

(The Soviet Journal *Avtomatika i Telemekhanika* in English Translation)

■ This translation of a Soviet journal on automatic control is published as a service to American science and industry. It is sponsored by the Instrument Society of America under a grant in aid from the National Science Foundation, continuing a program initiated by the Massachusetts Institute of Technology.

SOVIET INSTRUMENTATION AND CONTROL TRANSLATION SERIES

Instrument Society of America

Dr. Ralph H. Tripp
President

J. Johnston, Jr.
Past President

Philip A. Sprague
President-elect-Secretary

Henry J. Noebels
Dept. Vice President

E. A. Adler
Dept. Vice President

Adelbert Carpenter
Dept. Vice President

Nathan Cohn
Dept. Vice President

Francis S. Hoag
Dept. Vice President-elect

John C. Koch
Treasurer

Nelson Gildersleeve
Dist. I Vice President

H. Kirk Fallin
Dist. II Vice President

John R. Mahoney
Dist. III Vice President

F. R. Gilmer
Dist. IV Vice President

Milton M. McMillen
Dist. V Vice President

Otto J. Lessa
Dist. VI Vice President

J. Howard Park, III
Dist. VII Vice President

Roy Horton
Dist. VIII Vice President

Robert C. Mann
Dist. IX Vice President

Kenneth S. Vriesen
Dist. X Vice President

John J. McDonald
Dist. XI Vice President

International Headquarters

William H. Kushnick
Executive Director

Charles W. Covey
Editor, ISA Journal

Herbert S. Kindler
Director, Tech. & Educ. Services

Ralph M. Stotsenburg
Director, Promotional Services

ISA Publications Committee

Charles O. Badgett, *Chairman*

Jere E. Brophy, George A. Larsen, Joshua Stern, Dr. Enoch J. Durbin, Thomas G. MacAnespie, Frank S. Swa[illegible], Prof. Richard W. Jones, John E. Read, Richard A. Ter[illegible]

Translations Advisory Board of the Publications Committee

Jere E. Brophy, *Chairman*

T. J. Higgins, S. G. Eskin, G. Werbizky

■ This translation of the Soviet Journal *Avtomatika i Telemekhanika* is published and distributed at nominal subscription rates under a grant in aid to the Instrument Society of America from the National Science Foundation. This translated journal, and others in the Series (see back cover), will enable American scientists and engineers to be informed of work in the fields of instrumentation, measurement techniques, and automatic control reported in the Soviet Union.

The original Russian articles are translated by competent technical personnel. The translations are on a cover-to-cover basis and the Instrument Society of America and its translators propose to translate faithfully all of the scientific material in *Avtomatika i Telemekhanika*, permitting readers to appraise for themselves the scope, status, and importance of the Soviet work. All views expressed in the translated material are intended to be those of the original authors and not those of the translators nor the Instrument Society of America.

Publication of *Avtomatika i Telemekhanika* in English translation started under the present auspices in April, 1958, with Russian Vol. 18, No. 1 of January, 1957. The program has been continued with the translation and printing of the 1958-1961 issues.

Transliteration of the names of Russian authors follows the system known as the British Standard. This system has recently achieved wide adoption in the United Kingdom, and is currently being adopted by a large number of scientific journals in the United States.

Readers are invited to submit to the Instrument Society of America comments on the quality of the translations and the content of the articles. Pertinent correspondence will be published in the Society's monthly publication, the ISA JOURNAL. Space will also be made available in the ISA JOURNAL for such replies as may be received from Russian authors to comments or questions by the readers.

1961 Volume 22 Subscription Prices:

Per year (12 issues), starting with Vol. 22, No. 1

General: *United States and Canada* $35.00
Elsewhere 38.00

Libraries of nonprofit academic institutions:
United States and Canada $17.50
Elsewhere 20.50

Single issues to everyone, each . . . . $ 6.00

1957 Volume 18, 1958 Volume 19, 1959 Volume 20, and 1960 Volume 21 issues also available. Prices upon request.
See back cover for combined subscription to entire Series.

Subscriptions and requests for information on back issues should be addressed to the:

Instrument Society of America
530 William Penn Place, Pittsburgh 19, Penna.

Translated and printed by Consultants Bureau Enterprises, Inc.

Copyright © 1961 by the Instrument Society of America

Automation and Remote Control

A translation of Avtomatika i Telemekhanika, a publication of the Academy of Sciences of the USSR

EDITORIAL BOARD OF AVTOMATIKA I TELEMEKHANIKA

D. I. Ageikin
M. A. Aizerman
A. B. Chelyustkin (Assoc. Editor)
E. G. Dudnikov
N. Ya. Festa
V. A. Ilin
A. G. Iosuf'yan
V. V. Karibskii
A. V. Khramoi
B. Ya. Kogan
V. S. Kulebakin
S. A. Lebedev
A. Ya. Lerner
A. M. Letov (Assoc. Editor)
V. S. Malov
B. N. Petrov
Yu. P. Portnov-Sokolov
B. S. Sotskov
A. A. Tal' (Corresp. Secretary)
V. A. Trapeznikov (Editor in Chief)
Ya. Z. Tsypkin
G. M. Ulanov
A. A. Voronov
S. V. Yablonskii

Vol. 22, No. 1 — Russian Original Dated January, 1961 — August, 1961

CONTENTS

THE THEORY OF DUAL CONTROL. III

A. A. Fel'dbaum

Translated from Avtomatika i Telemekhanika, Vol. 22, No. 1,
pp. 3-16, January, 1961
Original article submitted April 18, 1960

We shall study two examples of the synthesis of an optimal system for dual control. In the first example, we study a system of automatic stabilization with a linear object in which the performance criterion is the mean-square error. In the second example, we study an automatic system which searches for the minimal output value for an object that has a parabolic characteristic. The analysis is performed on the basis of the theory presented in the preceding papers [1,2] in this series.

Remarks on the Synthesis of Optimal Dual-Control Systems

In [1,2] we formulated the statement of the problem for the case of synthesizing optimal dual-control systems, and presented a general method for solving this problem. In the analysis below we study examples of the application of the general theory to individual problems.

The method of solution consists of formulating a series of functions γ_{n-k} ($k = 0, 1, \ldots, n$), and finding those values u^*_{n-k} of the control input which minimize the functions γ_{n-k}. Assume the minimal value of the function γ_{n-k} with respect to u_{n-k} is denoted by γ^*_{n-k}. The function γ_{n-k-1} is formed from γ^*_{n-k} by integrating the output value y_{n-k} of the object which is measured by the control unit, and by adding the function α_{n-k-1}. Thus, the scheme for the basic computations involves a series of alternate integrations and minimizations which is clearly depicted in Fig. 1.* The results of the computations are the optimal inputs u^*_s.

In order to find the optimal control input u^*_0 at the initial instant $t = 0$, it is necessary to pass through the entire chain of computations shown in Fig. 1. At the instant $t = s$ the volume of the computations required for determining u^*_s is reduced, since it is necessary to perform only a portion of the chain of computations shown in Fig. 1 – from the end of u^*_n to the determination of u^*_s. The latter quantity is determined as a function of the vectors $\mathbf{u}_{s-1}$, $\mathbf{x}^*_s$, and $\mathbf{y}_{s-1}$; i.e., it is determined as a function of all the preceding values of the quantities at the input and output of the control section of the system. These values are memorized in the control section.

Depending on the complexity of the computations performed in the chain depicted in Fig. 1, the following cases may be encountered:

a. The entire chain of minimizations and averagings can be performed (either exactly or approximately) analytically. In that case, it is possible to compute the functions

$$u^*_s = u^*_s(\mathbf{u}_{s-1}, \mathbf{x}^*_s, \mathbf{y}_{s-1}) \qquad (s = 0, 1, \ldots, n) \tag{1}$$

in advance and to introduce the ready control law in the form of (1) into the control section A of the system. This case is treated in the first example in this paper. However, such a possibility is encountered only in the simplest examples.

*Figure 1 indicates integration with respect to x_i, since, in the subsequent examples, $h_s = 0$ and $y_i = x_i$.

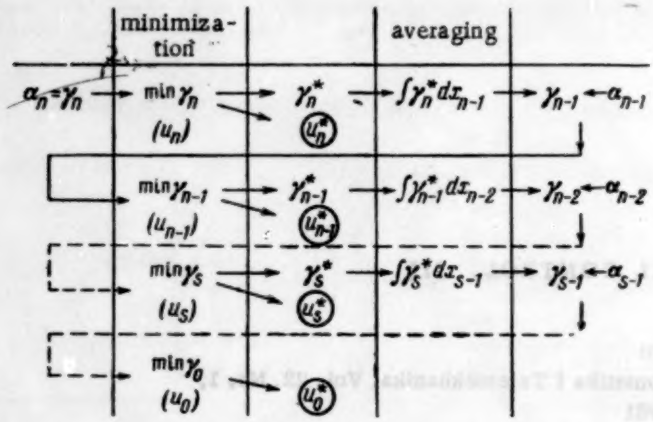


Fig. 1.

b. It is impossible to perform the entire chain of computations in advance, and to obtain ready formulas of the type (1). This is the more general case. Such a problem is treated in the second example. In this case, the theory provides the possibility of achieving the synthesis of an optimal dual-control system. Here we design a control section A for the system, such that it automatically performs (either completely or partially) the computations shown in Fig. 1.

The Synthesis of the Optimal Control Section for the Automatic Stabilization System

The block diagram of the system is shown in Fig. 2. The output quantity x of the controlled object B is applied to the input of the control section A of the system. The standard input x* (which is assumed constant), is applied to the other input of the section. The output quantity u from the section A of the system consists of the control input, which is applied to the input of the object B, together with a random constant bias μ and a random noise g. The values g_s of the noise at the instants t = s are a series of independent random quantities with the same distribution density $q(g_s)$. The a priori distribution density $P(\mu) = P_0(\mu)$ is specified by the random quantity μ. The object B has a linear characteristic of the following form:

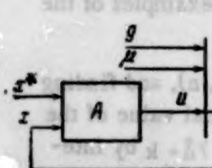


Fig. 2.

$$x_s = u_s + g_s + \mu. \tag{2}$$

The specific loss function is determined by the expression

$$W_s = (x_s - x^*)^2. \tag{3}$$

The over-all loss function is written as

$$W = \sum_{s=0}^{n} W_s = \sum_{s=0}^{n} (x_s - x^*)^2, \tag{4}$$

where the number of cycles n is specified. It is necessary to synthesize a control section A, such that the risk R (the mathematical expectation of the function W) is minimal.

In accordance with the general theory, it is required to find the probability density $P(y_i \mid \mu, i, u_i)$. Since, in this problem, y = x, μ is a scalar quantity, and the specified probability density is independent of i, we shall write it as $P(x_i \mid \mu, u_i)$. It is obvious that, in this problem,

$$P(x_i \mid \mu, u_i) = q(x_i - u_i - \mu). \tag{5}$$

In order to determine the optimal strategy, it is necessary to find the function $\alpha_k (0 \le k \le n)$, which is given by the expression

$$\alpha_k = \int\limits_{\Omega(x_k, \mu)} W_k P_0(\mu) \prod_{i=0}^{k} P(x_i | \mu, u_i)\, d\Omega(x_k, \mu) =$$

$$= \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} (x^* - x_k)^2 P_0(\mu) \prod_{i=0}^{k} q(x_i - u_i - \mu)\, dx_k\, d\mu. \tag{6}$$

We shall study the case where $P_0(\mu)$ and $q(g_s)$ are normal distributions with zero average values. Then

$$P_0(\mu) = \frac{1}{\sigma_\mu \sqrt{2\pi}} \exp\left\{-\frac{\mu^2}{2\sigma_\mu^2}\right\}, \qquad q(g_s) = \frac{1}{\sigma_g \sqrt{2\pi}} \exp\left\{-\frac{g_s^2}{2\sigma_g^2}\right\}. \tag{7}$$

From (6) and (7) we find

$$\alpha_k = \frac{1}{(\sigma_g)^{k+1} \sigma_\mu (2\pi)^{\frac{k}{2}+1}} \int\limits_{\mu=-\infty}^{+\infty} \exp\left\{-\left[\frac{\mu^2}{2\sigma_\mu^2} + \sum_{i=0}^{k-1} \frac{(x_i - u_i - \mu)^2}{2\sigma_g^2}\right]\right\} \times$$

$$\times \left[\int\limits_{x_k=-\infty}^{+\infty} (x^* - x_k)^2 \exp\left\{-\frac{(x_k - u_k - \mu)^2}{2\sigma_g^2}\right\} dx_k\right] d\mu. \tag{8}$$

We shall make use of the formula ([3], p. 188)

$$\int\limits_{-\infty}^{+\infty} e^{-px^2+2qx} x^{a+1}\, dx = \frac{1}{2^a p} \sqrt{\frac{\pi}{p}} \frac{d^a \left(q e^{\frac{q^2}{p}}\right)}{dq^a}, \tag{9}$$

where $\underline{a}$ is a positive integer, and $\underline{p}$ and $\underline{q}$ are rational.

For $a = 1$, we obtain

$$\int\limits_{-\infty}^{+\infty} x^2 e^{-px^2+2qx}\, dx = \frac{1}{2p} \sqrt{\frac{\pi}{p}}\, e^{\frac{q^2}{p}} \left(1 + \frac{2q^2}{p}\right). \tag{10}$$

We shall find the integral I_k in the square brackets of formula (8):

$$I_k = \int\limits_{-\infty}^{+\infty} (x^* - x_k)^2 \exp\left\{-\frac{(x_k - u_k - \mu)^2}{2\sigma_g^2}\right\} dx_k = \exp\left\{-\frac{l_k^2}{2\sigma_g^2}\right\} \int\limits_{-\infty}^{+\infty} z^2 \exp\left\{-\frac{z^2}{2\sigma_g^2} + 2\left(-\frac{l_k}{2\sigma_g^2}\right) z\right\} dz, \tag{11}$$

where

$$x_k - x^* = z, \qquad x^* - u_k - \mu = l_k. \tag{12}$$

Making use of formula (10), we find

$$I_k = \sigma_g^3 \sqrt{2\pi}\, [1 + (x^* - u_k - \mu)^2] \tag{13}$$

from (11) and (12).

Then

$$\alpha_k = \frac{1}{\sigma_g^{k-2} \sigma_\mu (2\pi)^{\frac{k+1}{2}}} \left[\int\limits_{-\infty}^{+\infty} \exp\left\{-\left[\frac{\mu^2}{2\sigma_\mu^2} + \sum_{i=0}^{k-1} \frac{(x_i - u_i)^2 - 2\mu(x_i - u_i) + \mu^2}{2\sigma_g^2}\right]\right\} d\mu + \right.$$

$$\left. + \int\limits_{-\infty}^{+\infty} \exp\left\{-\left[\frac{\mu^2}{2\sigma_\mu^2} + \sum_{i=0}^{k-1} \frac{(x_i - u_i)^2 - 2\mu(x_i - u_i) + \mu^2}{2\sigma_g^2}\right]\right\} (x^* - u_k - \mu)^2\, d\mu\right]. \tag{14}$$

The first of the integrals in the square brackets of (14) can be found if we use the formula ([3], p.185)

$$\int_{-\infty}^{+\infty} e^{-px^2 \pm qx}\,dx = e^{\frac{q^2}{4p}}\sqrt{\frac{\pi}{p}}. \tag{15}$$

The second integral in the square brackets can be determined by assuming $\mu + u_k - x^* = z$. Finally, the formula for α_k is obtained in finite form.

We shall assume

$$u_k - x^* = w_k, \qquad \sum_{i=0}^{k-1}(x_i - u_i) = \Sigma_{k-1},$$

$$\varepsilon_k = \frac{1}{2\sigma_\mu^2} + \frac{k}{2\sigma_g^2}, \qquad \sum_{i=0}^{k-1}(x_i - u_i)^2 = \theta_{k-1}, \tag{16}$$

$$a_k = \frac{1}{\sqrt{2}\sigma_g^{k-2}\sigma_\mu (2\pi)^{\frac{k-1}{2}}\sqrt{\varepsilon_k}}, \qquad b_k = \frac{\sqrt{\pi}}{2\sigma_g^{k-2}\sigma_\mu (2\pi)^{\frac{k+1}{2}}(\varepsilon_k)^{\frac{3}{2}}}.$$

Then

$$\alpha_k = \left\{a_k + b_k + \frac{2b_k}{\varepsilon_k}\left(w_k\varepsilon_k + \frac{\Sigma_{k-1}}{2\sigma_g^2}\right)^2\right\}\exp\left\{-\frac{\theta_{k-1}}{2\sigma_g^2} + \frac{\Sigma_{k-1}^2}{4\sigma_g^4\varepsilon_k}\right\}. \tag{17}$$

The quantity u_k, which is contained in w_k, is present in only one of the multipliers, and is not included in the exponential term. This fact greatly simplifies the minimization of α_k with respect to u_k.

We shall now perform the series of computations indicated in Fig. 1. In formula (17) we shall write k = n and find the minimum of the function $\gamma_n = \alpha_n$ with respect to u_n. It is obvious that the condition governing the minimum is written as

$$w_n\varepsilon_n + \frac{\Sigma_{n-1}}{2\sigma_g^2} = 0, \tag{18}$$

whence we find the optimal value u_n^* for the control input:

$$u_n^* = x^* - \frac{\sum_{i=0}^{n-1}(x_i - u_i)}{2\sigma_g^2\varepsilon_n}. \tag{19}$$

Then we find

$$\gamma_n^* = \alpha_n^* = (a_n + b_n)\exp\left\{-\frac{\theta_{n-1}}{2\sigma_g^2} + \frac{\Sigma_{n-1}^2}{4\sigma_g^4\varepsilon_n}\right\}. \tag{20}$$

After that we find the function γ_{n-1} (cf. [2]):

$$\gamma_{n-1} = \alpha_{n-1} + \int_{x_{n-1}=-\infty}^{+\infty} \alpha_n^*\,dx_{n-1} =$$

$$= \left\{a_{n-1} + b_{n-1} + \frac{2b_{n-1}}{\varepsilon_{n-1}}\left[w_{n-1}\varepsilon_{n-1} + \frac{\Sigma_{n-2}}{2\sigma_g^2}\right]^2\right\}\exp\left\{-\frac{\theta_{n-2}}{2\sigma_g^2} + \frac{\Sigma_{n-2}^2}{4\sigma_g^4\varepsilon_{n-1}}\right\} + \tag{21}$$

$$+ \int_{-\infty}^{+\infty}(a_n + b_n)\exp\left\{-\frac{(x_{n-1} - u_{n-1})^2 + \theta_{n-2}}{2\sigma_g^2} + \frac{[\Sigma_{n-2} + (x_{n-1} - u_{n-1})]^2}{4\sigma_g^4\varepsilon_n}\right\}dx_{n-1}.$$

We denote the integral in this formula by J_{n-1}. Using formula (15), we determine this integral:

$$J_{n-1} = (a_n + b_n)\sqrt{\frac{\pi}{\left(\frac{1}{2\sigma_g^2} - \frac{1}{4\sigma_g^4 \varepsilon_n}\right)}} \times$$
$$\times \exp\left\{\frac{\theta_{n-2}}{2\sigma_g^2} + \frac{\Sigma_{n-2}^2}{4\sigma_g^4\varepsilon_n} + \frac{\Sigma_{n-2}^2}{16\sigma_g^8\varepsilon_n^2\left(\frac{1}{2\sigma_g^2} - \frac{1}{4\sigma_g^4\varepsilon_n}\right)}\right\}. \quad (22)$$

In this formula, we encounter the quantity

$$\frac{1}{2\sigma_g^2} - \frac{1}{4\sigma_g^4\varepsilon_n} = \frac{1}{2\sigma_g^2} - \frac{1}{\frac{2\sigma_g^4}{\sigma_\mu^2} + 2n\sigma_g^2}.$$

This quantity is positive, since $n \geq 1$ and $\frac{2\sigma_g^4}{\sigma_\mu^2} > 0$. It should be noted that J_{n-1} is independent of u_{n-1}. Therefore, in formula (21), only the first term α_{n-1} depends on u_{n-1}; this facilitates minimization of γ_{n-1} with respect to u_{n-1}.

In this case, only α_{n-1} depends on u_{n-1}, and $\int \gamma_{n-1}^* \, \alpha x_{n-1}$ is independent of u_{n-1}; this means that only the operational risk for $t = n - 1$ depends on u_{n-1}, whereas the adaptive risk is independent of the quantity u_{n-1}, even though it is nonzero. Thus, the system under study is neutral [2], and in it the process of adapting to the object proceeds identically for any value of u_{n-1}, and, in general, for u_s ($0 \leq s < n$).

Minimizing α_{n-1} with respect to u_{n-1}, we find the condition governing the minimum in the same form as Eq. (18):

$$w_{n-1}\,\varepsilon_{n-1} + \frac{\Sigma_{n-2}}{2\sigma_g^2} = 0, \quad (23)$$

whence it follows that

$$u_{n-1}^* = x^* - \frac{\sum_{i=0}^{n-2}(x_i - u_i)}{2\sigma_g^2\varepsilon_{n-1}}. \quad (24)$$

Reasoning analogously, we find that, in general, γ_s^* is independent of w_s, and depends only on the difference $x_{s-1} - u_{s-1}$. Therefore, the substitution $z = x_{s-1} - u_{s-1}$ liquidates dependence on u_{s-1} in the integral J_{s-1}. Therefore, for any $s - 1$, only α_{s-1} depends on u_{s-1}, and the formula for determining u_s^* is written as

$$u_s^* = x^* - \frac{\sum_{i=0}^{s-1}(x_i - u_i)}{2\sigma_g^2\varepsilon_s}. \quad (25)$$

Substituting the value of ϵ_s from (16), we find the optimal control law in the following form:

$$u_s^* = x^* - \frac{\sum_{i=0}^{s-1}(x_i - u_i)}{s + \left(\frac{\sigma_g}{\sigma_\mu}\right)^2} \qquad (s = 0, 1, \ldots, n). \quad (26)$$

This formula has a simple meaning. If the noise g_s were to be equal to zero, then (Fig. 2) we would need only two values (namely x_i and u_i) for determining the bias μ; here $\mu = x_i - u_i$. It is obvious that, after this measurement, the value of u_{i+1}, which must be introduced, is equal to $x^* - \mu = x^* - (x_i - u_i)$. If the noise g_s exists, then, for large s, it is natural to determine μ in the form of the arithmetic mean of all the measurements of $(x_i - u_i)$, i.e., $\mu_{av} \approx \frac{\sum_{i=0}^{s-1}(x_i - u_i)}{s}$. Then we introduce the value

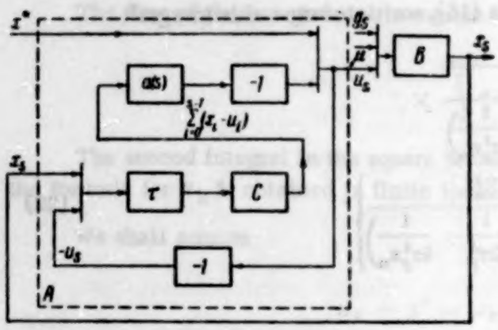


Fig. 3.

$$u_s = x^* - \mu_{av} = x^* - \frac{\sum_{i=0}^{s-1} (x_i - u_i)}{s}. \tag{27}$$

The exact formula (26) differs from formula (27), which is derived from simple considerations, only for small s. For sufficiently large s, the two formulas coincide.

This example is the simplest one, and is illustrative in nature, since the measurement of the difference $(x_i - u_i)$ makes it possible to determine the value of $\mu + g_s$, and then to determine the bias μ as if it had been measured directly with the error g_s. The same result would be obtained in the problem under study in the case of a nonlinear, mutually single-valued characteristic for the object. However, even in linear systems, the synthesis of an optimal control section A is appreciably complicated when two coordinates μ_1 and μ_2 of the vector $\boldsymbol{\mu}$ are present, if the characteristic of the object has, for example, the form

$$x_s = \mu_2 (u_s + \mu_1 + g_s). \tag{28}$$

The general theory cited in [2] makes it possible to solve this problem in principle.

From the optimal control law (26) it is not difficult to formulate the block diagram of the control system (cf. Fig. 3). The sign (-1) denotes an inverter (i.e., a sign-changing block). The letter τ denotes a block that produces a delay of one cycle, and the letter C denotes an adder with storage. The result of the summing (it is assumed that n is not so great that the adder overflows) is the quantity $\sum_{i=0}^{s-1} (x_i - u_i)$, which is subjected to multiplication by $\left[s + \left(\frac{\sigma_g}{\sigma_\mu}\right)^2\right]^{-1}$ in the block a(s), and then is subjected to sign inversion.

It is of definite interest to extrapolate this scheme to the case of a continuous system. We shall assume that the duration of a cycle is $\Delta t = t/s$, and shall make s go to infinity while maintaining $t = \text{const}$. Then $\Delta t \to 0$. Assume that we have the problem of synthesizing a control system for the case where the noise $g(t)$ is stationary white noise with the spectral density S_0. However, white noise is not the limit of a series of independent random quantities with finite dispersion, since its dispersion is infinitely great:

$$\sigma^2 = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S_0 \, d\omega = \infty.$$

Therefore, it is possible to obtain white noise in the limit for $\Delta t \to 0$ from a series of quantities g_s only in the case where the dispersion of these quantities is correspondingly increased. Assume the correlation function for the noise is written as

$$\begin{aligned} K_g(\tau) &= \frac{S_0}{\Delta t}\left(1 - \frac{|\tau|}{\Delta t}\right) = \sigma_g^2 \left(1 - \frac{|\tau|}{\Delta t}\right) \quad \text{for} \quad -\Delta t \leqslant \tau \leqslant \Delta t; \\ K_g(\tau) &= 0 \quad \text{for} \quad |\tau| > \Delta t. \end{aligned} \tag{29}$$

Then the spectral density is

$$S_g(\omega) = 2 \int_0^{\Delta t} K_g(\tau) \cos \omega\tau \, d\tau =$$

$$= \frac{2\sigma_g^2}{\omega} \sin \omega\Delta t - \frac{2\sigma_g^2}{\omega^2 \Delta t} [\omega\Delta t \sin \omega\Delta t + \cos \omega\Delta t - 1]. \tag{30}$$

For $\Delta t \to 0$, the function $S_g(\omega)$ tends toward $\sigma_g^2 \Delta t$:

$$\lim_{\Delta t \to 0} S_g(\omega) = \lim_{\Delta t \to 0} \sigma_g^2 \Delta t = \lim_{\Delta t \to 0} \frac{S_0}{\Delta t} \Delta t = S_0. \tag{31}$$

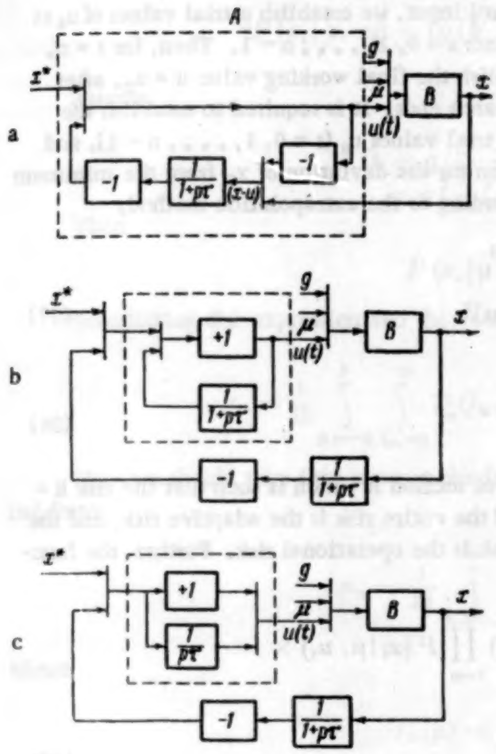


Fig. 4.

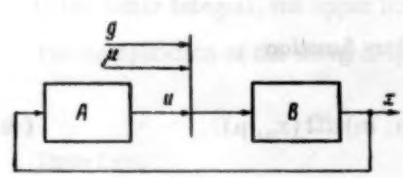


Fig. 5.

We shall rewrite formula (26) as follows:

$$u_s^* = x^* - \frac{\sum_{i=0}^{s-1}(x_i - u_i)\,\Delta t}{s\Delta t + \frac{\sigma_g^2 \Delta t}{\sigma_\mu^2}}. \qquad (32)$$

For $\Delta t \to 0$ the quantities in this formula tend toward the following limits:

$$\sigma_g^2 \Delta t \to S_0, \qquad \sum_{i=0}^{s-1}(x_i - u_i)\,\Delta t \to \int_0^t (x - u)\,dt. \quad (33)$$

Since $s\Delta t = t$, it follows that, in the limit, we obtain the following optimal control law from (32):

$$u^*(t) = x^* - \frac{\int_0^t (x - u)\,dt}{t + \frac{S_0}{\sigma_\mu^2}} = x^* - \frac{\int_0^t (x - u)\,dt}{t + a}, \quad (34)$$

where $a = S_0/\sigma_\mu^2$. The second term in this expression can be obtained at the output of a filter with the transient response

$$u_{\text{out}} = \frac{1}{t+a}\int_0^t u_{\text{in}}\,dt = \frac{t}{t+a}, \qquad (35)$$

if we apply the quantity $u_{in} = x - u$ to its input.* This transient response is not an exponential. The transfer function K(p) for such a filter is expressed in terms of an integral exponential function, and cannot be represented in the form of a finite sum of fractionally rational functions. However, it is possible approximately to replace the filter having the transient function (35) with an inertial section having the time constant $T \approx 1.1\ a$. Then a system which is close to optimal will be of the form illustrated in Fig. 4a. Converting this system to an equivalent one, we successively obtain the systems shown in Figs. 4b and 4c. The latter variant is a section which introduces the integral into the channel of the system, and has an inertial section in the feedback loop. It is of interest to compare these systems with similar ones which were obtained on the basis of completely different concepts by I. B. Reswick [4] and O. T. M. Smith [5].

The Synthesis of the Optimal Control Section for an Automatic Search System

We shall study an example of applying the theory of dual control for synthesizing an optimal automatic search system. The block diagram for the system is shown in Fig. 5. Assume that the characteristic for the object B is written as

$$x_s = (u_s + g_s + \mu)^2. \qquad (36)$$

It is required that the value of $\underline{x}$ prove to be as close as possible to the minimum as a result of automatic search. The constant bias μ at the object's input is not known a priori, and is a random quantity with a specified a priori probability density $P(\mu) = P_0(\mu)$. The noise g_s, just as in the preceding example, consists of a series of independent random quantities with a zero average value and the same probability density $P(g_s) = q(g_s)$.

*In determining the transient function, it was assumed that $u_{in} = 1 = \text{const}$.

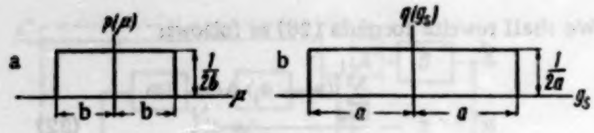


Fig. 6.

In order to determine the required value of the control input, we establish $\underline{n}$ trial values of u_s at the instants $s = 0, 1, \ldots, n-1$. Then, for $s = n$, we establish the final working value $u = u_n$, after which search ends. It is required to establish the series of trial values u_s $(s = 0, 1, \ldots, n-1)$, and the working value u_n in such a way that the mathematical expectation for the deviation of x_n from the minimum is minimal. Such a search will be an optimal search performed according to the extrapolation method.

The loss function for the problem under study will be written as

$$W_s = \vartheta_s x_s = \vartheta_s (u_s + g_s + \mu)^2, \tag{37}$$

where

$$\vartheta_s = \begin{cases} 0, & s < n, \\ 1, & s = n. \end{cases} \tag{38}$$

The problem consists of determining an algorithm for the control section A which is such that the risk $R = R_n = M\{W_n\}$ will be minimal. Therefore, for any $\underline{s}$th cycle $(s < n)$ the entire risk is the adaptive risk, and the operational risk is equal to zero. Moreover, for $s = n$, the over-all risk is the operational risk. Further, the function α_s is equal to zero for $s < n$. The quantity

$$R = R_n = \int\limits_{\Omega(x_n,\, \mu,\, u_n,\, x_{n-1})} x_n P(x_n \mid \mu, u_n) P(\mu) \prod_{i=0}^{n-1} P(x_i \mid \mu, u_i) \times \\ \times \prod_{i=0}^{n} \Gamma_i d\Omega(x_n, \mu, u_n, x_{n-1}) \tag{39}$$

is to be minimized.

In accordance with general theory, it is necessary to study the auxiliary function

$$\gamma_n = \alpha_n = \int\limits_{\Omega(x_n,\, \mu)} x_n P(x_n \mid \mu, u_n) P(\mu) \prod_{i=0}^{n-1} P(x_i \mid \mu, u_i)\, d\Omega(x_n, \mu). \tag{40}$$

If the quantities μ and g_s are normally distributed, then it follows that, from Eq. (40), it is possible to obtain the function α_n in finite form by integration. However, even in this case, only the first minimization (with respect to u_n) is not beset with difficulty. The subsequent integrations and minimizations already cannot be performed analytically. For other distribution laws, $P(\mu)$ and $q(g_s)$, the situation is complicated even at the first stage of the computations. In the analysis below we shall study the case where the densities $P(\mu)$ and $q(g_s)$ correspond to uniform distribution (Fig. 6). Assume also that the condition $b < a$ is satisfied. The meaning of the parameters $\underline{a}$ and $\underline{b}$ is clear from Fig. 6. Assume that

$$\eta_s = u_s + \mu + g_s. \tag{41}$$

Then $\eta_s^2 = x_s$, and $2\eta_s d\eta_s = dx_s$. From this it follows that*

$$\Pr(x' < x_s < x' + dx_s) = \\ = \Pr\left(\sqrt{x'} < \eta_s < \sqrt{x'} + \frac{dx_s}{2\sqrt{x'}}\right) + \Pr\left(-\sqrt{x'} < \eta_s < -\sqrt{x'} + \frac{dx_s}{2\sqrt{x'}}\right), \tag{42}$$

or

$$\Pr(x' < x_s < x' + dx_s) = \Pr\left(\sqrt{x'} - u_s - \mu < g_s < \sqrt{x'} - u_s - \mu + \frac{dx_s}{2\sqrt{x'}}\right) + \\ + \Pr\left(-\sqrt{x'} - u_s - \mu < g_s < -\sqrt{x'} - u_s - \mu + \frac{dx_s}{2\sqrt{x'}}\right). \tag{43}$$

* The letters Pr are the abbreviation for "probability."

Thus,

$$P(x_i \mid \mu, u_i) = [q(\sqrt{x_i} - u_i - \mu) + q(-\sqrt{x_i} - u_i - \mu)]\frac{1}{2\sqrt{x_i}}. \tag{44}$$

Assume

$$\begin{gathered}\sqrt{x_i} = \zeta_i,\\ Q_i = q(\zeta_i - u_i - \mu) + q(-\zeta_i - u_i - \mu) =\\ = q(-\zeta_i + u_i + \mu) + q(\zeta_i + u_i + \mu).\end{gathered} \tag{45}$$

Then

$$P(x_i \mid \mu, u_i) = \frac{1}{2\zeta_i} Q_i(\zeta_i, u_i + \mu). \tag{46}$$

Substituting this expression and the value of $P(\mu)$ into (40), we find

$$\alpha_n = \frac{1}{2b}\int_{\mu=-b}^{b}\int_{\zeta_n=0}^{\infty} \zeta_n^2 Q_n(\zeta_n, u_n + \mu)\prod_{i=0}^{n-1}\frac{1}{2\zeta_i} Q_i(\zeta_i, u_i + \mu)\, d\mu\, d\zeta_n. \tag{47}$$

Here we should take into account the fact that $dx_n = 2\zeta_n d\zeta_n$. Formula (47) can be written in the following form:

$$\alpha_n = \frac{1}{2b}\int_{\mu=-b}^{b} I_n(\mu)\prod_{i=0}^{n-1}\frac{1}{2\zeta_i} Q_i(\zeta_i, u_i + \mu)\, d\mu, \tag{48}$$

where

$$I_n(\mu) = \int_{\zeta_n=0}^{\infty} \zeta_n^2 Q_n(\zeta_n, u_n + \mu)\, d\zeta_n. \tag{49}$$

In the latter integral, the upper limit may be finite, since $Q_n = 0$ for $\zeta_n > a + 2b$.

The computation of the integral $I_n(\mu)$ yields the formula

$$I_n(\mu) = \frac{a^2}{3} + (u_n + \mu)^2. \tag{50}$$

Therefore,

$$\begin{aligned}\alpha_n = \frac{1}{2b}\left(\prod_{i=0}^{n-1}\frac{1}{2\zeta_i}\right)\Bigg[\frac{a^2}{3}\int_{-b}^{b}\prod_{i=0}^{n-1} Q_i(\zeta_i, u_i + \mu)\, d\mu +\\ + \int_{-b}^{b}(u_n + \mu)^2\prod_{i=1}^{n-1} Q_i(\zeta_i, u_i + \mu)\, d\mu\Bigg].\end{aligned} \tag{51}$$

We shall assume that

$$\begin{gathered}I_{0,n} = \int_{-b}^{b}\prod_{i=0}^{n-1} Q_i(\zeta_i, u_i + \mu)\, d\mu, \quad I_{1,n} = 2\int_{-b}^{b}\mu\prod_{i=0}^{n-1} Q_i(\zeta_i, u_i + \mu)\, d\mu,\\ I_{2,n} = \int_{-b}^{b}\mu^2\prod_{i=0}^{n-1} Q_i(\zeta_i, u_i + \mu)\, d\mu.\end{gathered} \tag{52}$$

Then formula (51) will become

$$\alpha_n = \frac{1}{2b}\left(\prod_{i=0}^{n-1}\frac{1}{2\zeta_i}\right)\left[\frac{a^2}{3} I_{0,n} + u_n^2 I_{0,n} + u_n I_{1,n} + I_{2,n}\right]. \tag{53}$$

Differentiating α_n with respect to u_n, we find the optimal value of u_n^*:

$$u_n^* = -\frac{I_{1,n}}{2I_{0,n}}. \tag{54}$$

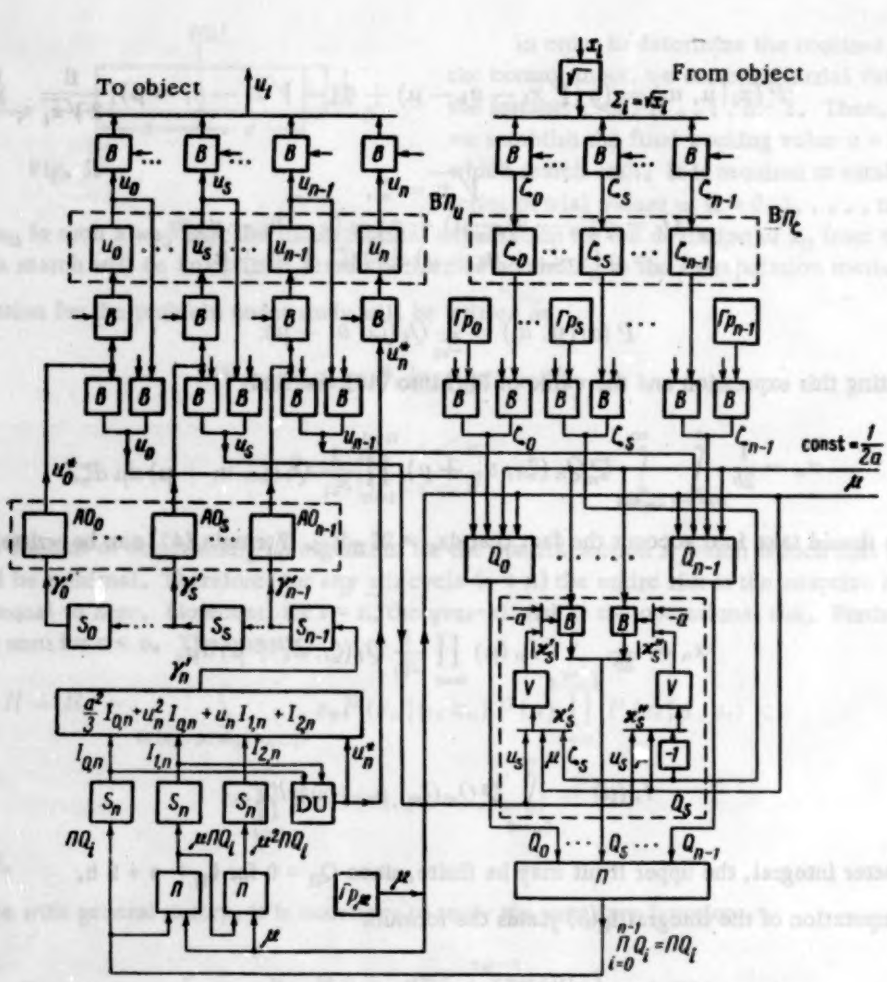


Fig. 7.

and the corresponding value of α_n^*:

$$\alpha_n^* = (\alpha_n)_{u_n = u_n^*} = \frac{1}{2b}\left(\prod_{i=0}^{n-1}\frac{1}{2\zeta_i}\right)\left[\frac{a^2}{3}I_{0,n} + I_{2,n} - \frac{I_{1,n}^2}{4I_{0,n}}\right]. \tag{55}$$

The expressions $I_{0,n}$, $I_{1,n}$, $I_{2,n}$ in the form of formulas are so complex that it is simpler to determine the integrals in the actual control section A. The further integrations and minimizations in formula (55) can also be performed only within A, since the formulas cannot be obtained in finite form. However, the over-all sequence of computations can be realized by a computer device whose block diagram is shown in Fig. 7. This is the block diagram for an optimal control system. The blocks in this system can, in principle, be either analog or digital, and the diagram proper can be replaced by its equivalent program on a universal digital computer.

In the block diagram shown in Fig. 7 we have used the following notation: B are rectifiers which are controlled by a control and synchronization block not shown in the figure; Γp are generators which produce "triangular" waves at various frequencies; V are the modulus blocks for which the output quantity is equal to the absolute value of the input quantity; S_i are integrating blocks (a more detailed diagram of the integrating block is shown in Fig. 8); AO_i are automatic optimizers with respect to one variable; Π are the product blocks; DU is a dividing unit; $BП_u$ is the memory block for the u_i values; $BП_\zeta$ is the memory block for ζ_i values; √ is a block whose output quantity is $\zeta_i = \sqrt{x_i}$, where x_i is the input quantity.

The generator Γp_{μ} applies a voltage to the bus μ, which varies over the limits $-b \leq \mu \leq b$ according to a "triangular" curve; here $\left|\frac{d\mu}{dt}\right| = \text{const}$. The blocks $Q_0, Q_1, \ldots, Q_s, \ldots, Q_{n-1}$ form the corresponding functions in accordance with Eq. (45). As an example, Fig. 7 shows the block diagram for the section Q_s in more detail; here Q_s is the sum of the output signals from two rectifiers B. Each of these output signals is equal to $\frac{1}{2}a$ or zero, depending on the conditions

$$|\varkappa_s'| = |u_s + \mu + \zeta_s| \lessgtr a, \qquad |\varkappa_s''| = |u_s + \mu - \zeta_s| \gtrless a. \tag{56}$$

The terms $\underline{q}$ in expression (45) are obtained at the outputs of the rectifiers.

All of the functions Q_s ($s = 0, 1, \ldots, n-1$) are applied to the block Π, where the product $\prod_{i=0}^{n-1} Q_i = \Pi Q_i$ is produced. The resulting product, and also the products $\mu \Pi Q_i$ and $\mu^2 \Pi Q_i$, are applied to integrating blocks where integration with respect to μ is performed during a half-cycle of the voltage from the generator Γp_{μ}. The resulting values $I_{0,n}$, $I_{1,n}$, and $I_{2,n}$ are applied to the input of the block, which computes the function

$$\gamma_n' = \frac{a^2}{3} I_{0,n} + u_n^2 I_{0,n} + u_n I_{1,n} + I_{2,n} \tag{57}$$

[i.e., that factor in the function $\alpha_n = \gamma_n$ which depends on u_n in accordance with formula (53)]. At the output of the dividing unit DU, the normal value of u_n^* as computed according to Eq. (54) is determined; the quantity u_n^* is applied to the input of the block γ_n'. Therefore, at its output, we obtain the value $\gamma_n'^*$ corresponding to $u_n = u_n^*$. The function γ_n' is integrated in the block S_{n-1} during one half-cycle of the generator Γp_{n-1}, i.e., it is integrated with respect to the coordinate ζ_{n-1}, which, for the time being, arrives from the generator Γp_{n-1}. As a result, we obtain the function γ_{n-1}'. In order to integrate we must first multiply α_n^* [cf. (55)] by $2\zeta_{n-1} d\zeta_{n-1}$, and then integrate. Therefore, only the square bracket in formula (55) must participate in the integration. The multipliers $1/\zeta_i$ must not be taken into account.

The output signal γ_{n-1}' of the block S_{n-1} is transmitted to the input of the automatic optimizer AO_{n-1}, which chooses the value of u_{n-1}^* in such a way as to ensure the minimum for γ_{n-1}'. The next integration over a half cycle of the generator Γp_{n-2} in the integrator S_{n-2} makes it possible to obtain the function γ_{n-2}', etc. The chain of optimizers automatically selects the values of $u_0^*, u_1^*, \ldots, u_{n-1}^*$, which minimize the functions $\gamma_0', \gamma_1', \ldots, \gamma_{n-1}'$, respectively.

The value of u_0^* chosen in this manner is applied to the memory block $B\Pi_u$ and stored in the cell u_0. From this it is transmitted to the object at the instant $t = 0$. As soon as the value u_0 is memorized in the $B\Pi_u$ and is applied to the object, the computation of u_0 can be ceased, and from then on u_0 is applied to the block Q_0 only from the memory block. Analogously, at the instant following the memorization of the value u_1, the latter begins to be applied to the block Q_1 only from the memory block $B\Pi_u$, etc. Gradually, during the process of operation of the network, all of the cells u_i in the memory block are filled. Before the last cycle, the cell u_n is filled. In exactly the same way, all of the cells ζ_i in the memory block $B\Pi_{\zeta}$ are gradually filled. As soon as the value ζ_s has appeared at the output of the object at the instant $t = s$, it is stored in the $\underline{s}$th cell of the ζ_s memory block, and from that instant the block Q_s is subjected not to the output signal of the generator Γp_s, but to the constant value ζ_s which is measured by the block.

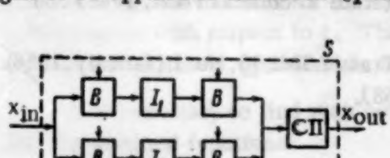


Fig. 8.

Figure 8 shows the block diagram of the integrator S. During the first half-cycle, the upper branch of the integrator I_1 operates (starting from zero) and accumulates the value of the integral. During the second half-cycle, the lower branch operates with an analogous integrator I_2, and during that time the accumulated reading is taken from I_1 and stored in the output memory cell CΠ. Thus, we obtain the required value of the integral at the output with a lag equal to a half-cycle of the corresponding generator.

The network in Fig. 7 is very complex. Its basic shortcoming, however, is not its complexity, but the necessity of an extremely rapid operation of a number of sections. The number of memory cells in the network is small; the number of blocks which perform the basic operations is also small, but the fundamental frequency

must be very great. Assume, for example, that $n = 3$. If the generator Γp_μ operates at a frequency of $f = 5$ Mc, then, during a half-cycle (i.e., during a period $T = 10^{-7}$ sec), the integrators S_n produce the values $I_{0,n}, I_{1,n}, I_{2,n}$. The quantity γ_n^* is applied to the integrator S_{n-1} with exactly the same lag. For integration with respect to ζ_{n-1}, it is required that no less than 10-20 values of γ_n^* correspond to a half-cycle of generator $\Gamma p_{n-1}(\Gamma p_2)$. Assuming that 20 half-cycles of Γp_μ must be used for one half-cycle of the generator Γp_2, we obtain a frequency of 250 kc for the latter. Thus, during one of its half-cycles (i.e., during a period $T_2 = 2 \cdot 10^{-6}$ sec), the value γ_2^* is produced. In order to compute $u_{n-1}^* = u_2^*$, it is necessary to pass over the entire scale $-b \le u_2 \le b$ while testing not less than 10 to 20 positions on the scale. Therefore, the computation of u_2^* requires a time equal to $T_2^* = 20T_2 = 40 \cdot 10^{-6}$ sec. Reasoning analogously, we find that integration for obtaining γ_1^* requires a time equal to $T_1 \approx 20T_2^* = 8 \cdot 10^{-4}$ sec, and for obtaining u_1^* it requires a time equal to $T_1^* \approx 20T_1 = 1.6 \cdot 10^{-2}$ sec. Integration for the purpose of forming γ_0^* requires a time $T_0 = 20T_1^* = 32 \cdot 10^{-2}$ sec, and a time $T_0^* = 20T_0 = 20 \cdot 0.32 = 6.4$ sec is required to obtain u_0^*. Thus, even the very high carrier frequency of 5 Mc (and at that for a relatively small value $n = 3$) cannot assure a short time for determining the first value u_0. Of course, the subsequent values u_i will be determined more rapidly. If we were to determine the value of u_n^* not according to formula (54) but by minimization of γ_n^*, then the time for determining u_0 would prove to be still greater.

This example demonstrates that, as far as possible, we should attempt an approximate solution of the problem involving the analytical determination of the values u_i^* since the computation of these values in a computer requires an extremely high speed of response. Modern digital machines cannot insure the required speed of response when complex problems are solved. This example also indicates the importance of papers in the field of designing ultrahigh-speed computers as far as the synthesis of optimal control systems is concerned.

However, we should keep in mind the fact that, for the principle of computer operation cited above, the operating time is of the order of a^n, where $a =$ const. Therefore, as $\underline{n}$ increases, the time increases extremely rapidly. If we perform the greater portion of the computations in advance, and feed already formulated functions into the computer, then the operation is accelerated. However, under these conditions, it is necessary to remember functions of $\underline{n}$ variables, and thus the volume of the memory will be equal to a number of the order of b^n, where $b =$ const. Only the approximation of functional relationships in the form of functions of a definite type with a comparatively small number of unknown parameters, combined with an acceleration of the operation of computers and the development of approximate methods for solving problems, will create the practical possibility of designing optimal dual-control devices in cases which involve any appreciable complexity.

LITERATURE CITED

1. A. A. Fel'dbaum, "Theory of dual control. I," Avtomatika i Telemekhanika 21, No. 9 (1960).
2. A. A. Fel'dbaum, "Theory of dual control. II," Avtomatika i Telemekhanika 21, No. 11 (1960).
3. I. M. Ryzhik, Tables of Integrals, Sums, Series, and Products [in Russian] (State Technical Press, 1948) 2nd edition.
4. I. B. Reswick, "Disturbance-response feedback. A new control concept," Trans. ASME 78, No. 1 (January, 1956).
5. O. T. M. Smith, Feedback Control Systems (McGraw-Hill, New York, 1958).

All abbreviations of periodicals in the above bibliography are letter-by-letter transliterations of the abbreviations as given in the original Russian journal. *Some or all of this periodical literature may well be available in English translation.* A complete list of the cover-to-cover English translations appears at the back of this issue.

OPTIMUM PROCESSES IN SYSTEMS WITH DISTRIBUTED PARAMETERS

A. G. Butkovskii

Translated from Avtomatika i Telemekhanika, Vol. 22, No. 1,
pp. 17-26, January, 1961
Original article submitted May 23, 1960

This article is devoted to the problem of the optimum control of systems whose motion is generally described by nonlinear integral relationships that correlate the command inputs with the output variables of the system to be controlled. It is shown how the obtained results can be applied to the solution of problems in the optimum control of systems described by partial differential equations in the case where the equation solution is expressed by an integral relationship.

Important results in the theory of optimum processes have been obtained by using nonclassical variation methods in investigating lumped-parameter systems that are described by ordinary differential equations [1,2]. However, there is a large number of problems where the systems to be controlled have distributed parameters. Such systems pertain, for instance, to a large number of production-line industrial processes, in particular the heating of metals in through-passage furnaces, the drying of strip and friable materials, continuous etching and deposition of coatings, agglomeration, distillation, etc. The problems in the optimum control of systems with distributed parameters are formulated in [3].

Statement of the Problem

Many problems in the optimum control of systems with distributed parameters can be solved on the basis of the results obtained in considering the following optimum control problem.

Let n coordinates of the vector $Q = Q(t) = [Q_1(t), \ldots, Q_n(t)]$ satisfy the equation

$$Q_i = Q_i(t) = \int_{t_0}^{t_1} K_i(t, \tau, u_1(\tau), \ldots, u_r(\tau))\, d\tau, \tag{1}$$

where

$$K_i(t, \tau, u_1, \ldots, u_r) \quad (i = 1, 2, \ldots, n)$$

are certain given functions of the arguments. The control functions $u_k = u_k(t)$ $(k = 1, 2, \ldots, r)$, are sectionally continuous with respect to t. The control vector $u = u(t) = [u_1(t), \ldots, u_r(t)]$ belongs to a closed region Ω at any instant of time t contained in the $t_0 \le t \le t_1$ interval.

It is necessary to find such control $u = u(t) \in \Omega$, $t_0 \le t \le t_1$ for which $Q(t_1) = Q_*$ (Q_* is the assigned vector), and the assigned functional

$$Q_0 = \int_{t_0}^{t_1} F(\tau, Q(\tau), u(\tau))\, d\tau \tag{2}$$

assumes the minimum possible value.

Optimum Regime Conditions

The necessary optimum regime condition which the control vector $u = u(t)$ $(t_0 \le t \le t_1)$ must satisfy can be formulated by the following maximum theorem.

Theorem 1. Let $u = u(t) \in \Omega$ $(t_0 \le t \le t_1)$ represent such control for which $Q(t_1) = Q_*$, where Q_* is the assigned vector. For the optimum control $u = u(t)$, and the corresponding trajectory $Q = Q(t)$, there should exist such a constant nonzero vector $c = (c_0, c_1, \ldots, c_n)$ that $c_0 \le 0$ and that, for almost all t, $t_0 \le t \le t_1$, the function

$$H = c_0 \int_{t_0}^{t_1} \sum_{i=1}^{n} \frac{\partial F(\tau, Q(\tau), u(\tau))}{\partial Q_i} K_i(\tau, t, u)\, d\tau +$$

$$+ c_0 F(t, Q(t), u) + \sum_{i=1}^{n} c_i K_i(t_1, t, u) \tag{3}$$

of the variable $\underline{u}$, $u \in \Omega$, attains its maximum at the point $u = u(t)$.

In certain problems, the point $Q(t_1)$ must belong to a certain multiplicity set M with an arbitrary dimension that does not exceed $n - 1$. This is the so-called problem with a mobile right-hand end.

If the point $Q(t_1) \in M$ is known, then, on the basis of the above theorem, optimum control, as applied to the problem with a mobile end, must satisfy the maximum condition. An additional condition determining the point $Q(t_1)$ in the multiplicity set M is the transversality condition [1]. Let $Q(t_1) \in M$ be the optimum trajectory end. We shall say that the transversality condition is satisfied if the $(c_1, \ldots, c_n)$ vector is normal to the plane S, which is tangent to the multiplicity set M at the $Q(t_1)$ point. The necessary condition for the optimum regime in the mobile right-hand end problem can now be formulated in the following manner.

<u>Theorem 2.</u> Let $u = u(t) \in \Omega$ $(t_0 \le t \le t_1)$, for which $Q(t_1) \in M$. For the optimum control $u = u(t)$, and the corresponding trajectory $Q = Q(t)$ in the mobile right-hand end problem, it is necessary that there exist a vector $c = (c_0, c_1, \ldots, c_n)$ different from zero, which will satisfy the conditions of theorem 1, and that, moreover, the transversality condition at the point $Q(t_1)$ be satisfied.

The proofs of theorems 1 and 2, which are based on the application of "acicular" variations, and on the properties of convex multiplicity sets [1], are given in Appendix I.

In certain problems, instead of the system (1), systems of the following form must be considered:

$$Q_i = Q_i(t) = \int_{t_0}^{t} K_i(t, \tau, u(\tau))\, d\tau \quad (i = 1, 2, \ldots, n). \tag{4}$$

In this, the condition $Q(t_1) = Q_*$ also must be satisfied for $t = t_1$, and the functional (2) must assume the minimum possible value. This problem represents a particular case of the previous problem if the $K_i(t, \tau, u)$ function in system (1) satisfies the condition

$$K_i(t, \tau, u(\tau)) \equiv 0 \quad (i = 1, 2, \ldots, n), \tag{5}$$

where

$$t \le \tau \le t_1.$$

In this case, theorems 1 and 2 are also valid and, because of (5), the function H will assume the following form:

$$H = c_0 \int_{t}^{t_1} \sum_{i=1}^{n} \frac{\partial F(\tau, Q(\tau), u(\tau))}{\partial Q_i} K_i(\tau, t, u)\, d\tau +$$

$$+ c_0 F(t, Q(t), u) + \sum_{i=1}^{n} c_i K_i(t_1, t, u). \tag{6}$$

Application of the Obtained Results

As an example of the utilization of the obtained results in solving problems in the optimum control of systems with distributed parameters, we shall consider two problems.

The first problem pertains to heat-exchange processes where it is necessary to secure, in the minimum time, a temperature distribution in a solid which would be close to the assigned distribution by means of an external temperature field, or to minimize in a certain time deviations (in any sense) of the temperature distribution from a certain assigned distribution.

Assume that the $q = q(x,t)$ function describes the temperature distribution in a solid in dependence on the space coordinate $\underline{x}$ $(0 \leq x \leq x_1)$ and the time $\underline{t}$ $(0 \leq t \leq t_1)$, where x_1 and t_1 are certain given values. Further, we shall assume that the heating equations are given by

$$\frac{\partial q}{\partial t} = a \frac{\partial^2 q}{\partial x^2} \quad (0 \leqslant t \leqslant t_1,\ 0 \leqslant x \leqslant x_1). \tag{7}$$

The boundary conditions are

$$\left.\frac{\partial q}{\partial x}\right|_{x=x_1} = \alpha\,[u(t) - q(x_1, t)], \tag{8}$$

$$\left.\frac{\partial q}{\partial x}\right|_{x=0} = 0. \tag{9}$$

The initial condition is

$$q(x, 0) = 0. \tag{10}$$

The optimum control problem consists in determining a control function $u = u(t)$, restricted by the conditions $N_1 \leq u(t) \leq N_2$ $(0 \leq t \leq t_1)$, for which the integral of the square of the solid temperature distribution deviation $q_1 = q(x,t_1)$ from a certain assigned constant temperature q_a assumes the least possible value at the instant of time t_1, i.e.,

$$I = \int_0^{x_1} [q_a - q(x, t_1)]^2\, dx = \min. \tag{11}$$

As is known from the theory of partial differential equations [4,5,6], the function $q = q(x,t)$ that satisfies the conditions (7)-(10) can be represented in the following form:

$$q = q(x, t) = \int_0^t \varphi(x, t, \tau)\, u(\tau)\, d\tau, \tag{12}$$

where φ is a known function of its arguments (the Green function method).

The minimization of functional (11) can be reduced (see Appendix II) to the minimization of the functional

$$Q_0 = \int_0^{t_1} [Q_1(\tau) - \gamma R(t_1, \tau)]\, u(\tau)\, d\tau \tag{13}$$

under the following conditions:

$$Q_1(t) = \int_0^{t_1} T(t, \tau)\, u(\tau)\, d\tau, \qquad Q_2 = \int_0^{t_1} d\tau, \tag{14}$$

where R and T are certain known functions, and γ is a number.

In order to find the extremal control function $u = u(t)$, we shall apply theorem 1.

The H function has the following form:

$$H = c_0 \int_0^{t_1} u(\tau)\, T(t, \tau)\, u(t) d\tau + c_0 [Q_1(t) - \gamma R(t_1, t)] u(t) + c_1 T(t, t_1) u(t) + c_2. \tag{15}$$

From the transversality condition (theorem 2), we have $c_1 = 0$; if we take into account that $c_0 \leq 0$, the function H has a maximum if the following condition is satisfied:

$$u(t) = \frac{1}{2}(N_1 + N_2) + \frac{1}{2}(N_2 - N_1) \times$$
$$\times \operatorname{sign}\left\{\gamma R(t_1, t) - \int_0^{t_1} [T(t, \tau) + T(\tau, t)]\, u(\tau)\, d\tau\right\}. \tag{16}$$

Condition (16) has the form of an integral equation for the u = u(t) function.

For the v = v(t) function, which is determined by the relation

$$u = \frac{1}{2}(N_2 + N_1) + \frac{1}{2}(N_2 - N_1)v, \tag{17}$$

this equation is equivalent to the equation

$$v(t) = \operatorname{sign}\left[a(t) + \int_0^{t_1} b(t, \tau)\, v(\tau)\, d\tau\right], \tag{18}$$

where $a(t)$ and $b(t,\tau)$ are certain known functions which are expressed in terms of the known functions in Eq. (16), while $|v(t)| \le 1$ $(0 \le t \le t_1)$.

We shall assume that the expression under the symbol "sign" in Eqs. (16) and (18) does not become identically equal to zero, and that there exists a solution v = v(t) which has a finite number of discontinuities in the $(0, t_1)$ interval. In this case, the following method can be used for solving Eq. (18). We shall first check whether the $v(t) \equiv 1$ function or the $v(t) \equiv -1$ function is the solution of this equation for $0 \le t \le t_1$. If not, we shall assume that the solution v = v(t) changes its sign only once at the point θ_1 $(0 < \theta_1 < t_1)$, and we shall calculate the expressions Φ_1 and Φ_2 under the symbol "sign" in Eq. (18) as functions of time t and the parameter θ_1. In this, $\Phi_1 = \Phi_1(t, \theta_1)$ is calculated under the assumption that v = 1 in the first interval, and $\Phi_2 = \Phi_2(t, \theta_1)$ is calculated under the assumption that v = −1 in the first interval. Then, if there exists a solution v = v(t) with a single sign reversal, at least one of the equations $\Phi_1(\theta_1, \theta_1) = 0$, $\Phi_2(\theta_1, \theta_1) = 0$ must have a real solution. If no such solution exists, the solution should be successively sought, first in the class of sectionally constant functions (equal to 1 or −1) with two sign changes, then with three, etc., by successively calculating the Φ_1 and Φ_2 functions in dependence on the parameters $\theta_1, \theta_2, \theta_3$, etc. Finally, at a certain mth stage, as a result of the assumptions made, there must exist a solution v = v(t) with sign changes at m points $\theta_1, \theta_2, \ldots, \theta_m$, i.e., there must be a real solution of at least one of the two systems of equations

$$\Phi_1(\theta_i, \theta_1, \ldots, \theta_m) = 0 \quad (i = 1, 2, \ldots, m),$$

or

$$\Phi_2(\theta_i, \theta_1, \ldots, \theta_m) = 0 \quad (i = 1, 2, \ldots, m) \tag{19}$$

$$(0 < \theta_1 < \theta_2 < \ldots < \theta_m < t_1).$$

Here, the $\Phi_1 = \Phi_1(t_1, \theta_1, \ldots, \theta_m)$ function denotes the expression under the symbol "sign" in Eq. (18) in dependence on time t $(0 \le t \le t_1)$ under the assumption that $v(t) = 1$ in the first interval for $0 \le t \le \theta_1$. The $\Phi_2 = \Phi_2(t, \theta_1, \ldots, \theta_m)$ function denotes the same under the assumption that v = −1 in the first interval for $0 \le t \le \theta_1$.

Let us now consider the optimum control problem as applied to the process of heat exchange between a stationary and a mobile medium. Such processes take place, for instance, in heating billets in through-passage furnaces.

Let q = q(x,y,t) describe the temperature distribution in an infinitely wide strip with the thickness x_1, which moves with the velocity $\nu = \nu(t) \ge 0$ in the y axis positive direction. The x axis is perpendicular to the plate plane, and the x coordinate varies within the $0 \le x \le x_1$ interval. The heating of the material through the plate upper surface takes place in the $0 \le y \le y_1$ interval.

The heating equations have the following form:

$$\frac{\partial q}{\partial t} = a\frac{\partial^2 q}{\partial x^2} - \nu\frac{\partial q}{\partial y} \quad (0 \leqslant x \leqslant x_1,\ 0 \leqslant y \leqslant y_1,\ 0 \leqslant t \leqslant t_1). \tag{20}$$

The boundary conditions are

$$\left.\frac{\partial q}{\partial x}\right|_{x=x_1} = \alpha\,[u(t) - q(x_1, y, t)], \quad (21)$$

$$\left.\frac{\partial q}{\partial x}\right|_{x=0} = 0, \quad (22)$$

$$q(x, 0, t) = 0. \quad (23)$$

The initial condition is

$$q(x, y, 0) = q_0(x, y). \quad (24)$$

Here, $\underline{a}$ and α are positive constant coefficients, and the thickness x_1 of the solid can be a function of

$$\eta = y - \int_0^t \nu(p)\,dp.$$

In this case, the optimum control problem consists in determining a control function $u = u(t)$, restricted by the $N_1 \leq u(t) \leq N_2$ condition, for which the integral of the square of the deviation of the solid average temperature, with respect to its cross section

$$\bar{q} = \bar{q}(y, t) = \int_0^{x_1} q(x, y, t)\,dx,$$

from the assigned temperature q_a, assumes the minimum possible value at the point $y = y_1$ in the time interval from 0 to t_1, i.e.,

$$I = \int_0^{t_1} [q_a - \bar{q}(y_1, t)]^2\,dt = \min. \quad (25)$$

By the nondegenerate substitution of variables

$$\xi = x, \quad \eta = y - \int_0^t \nu(p)\,dp, \quad \tau = t, \quad (26)$$

Equation (20) can be reduced to the one-dimensional heat-conduction equation, and the $q = q(x,y,t)$ function that satisfies conditions (20)-(24) can be represented in the following form [4,5,6]:

$$q = q(x, y, t) = \int_0^{x_1} g(x, t, \xi)\, q_0\Big(\xi, y - \int_0^t \nu(p)\,dp\Big)\,d\xi + \int_0^t \varphi(x, t, \tau)\, u\Big(y - \int_\tau^t \nu(p)\,dp, \tau\Big)\,d\tau. \quad (27)$$

Here, $\underline{g}$ and φ are certain known functions.

If we assume that t_1 satisfies the condition $\int_0^{t_1} \nu(p)\,dp \leqslant y_1$, then, by taking (27) into account, the average with respect to the temperature cross section $\bar{q} = \bar{q}(y,t)$ will be given by

$$\bar{q} = \bar{q}(y, t) = G(y, t) + \int_0^t k(t, \tau)\, u(\tau)\,d\tau. \quad (28)$$

Here, $G(y,t)$ is a known function representing the first term in (27), averaged with respect to $\underline{x}$, and $k(t,\tau)$ is the $\varphi(x,t,\tau)$ function, which is averaged with respect to $\underline{x}$.

By denoting $Q_1(t) = \int_0^t k(t, \tau) u(\tau)\, d\tau$ and substituting (28) in expression (25) for $y = y_1$, the problem will be reduced to the minimization of the functional

$$Q_0 = \int_0^{t_1} [\Psi(\tau) - Q_1(\tau)]^2\, d\tau \tag{29}$$

[where $\Psi(\tau)$ is a known function], under the condition that

$$Q_1(t) = \int_0^t k(t, \tau)\, u(\tau)\, d\tau, \qquad Q_2 = \int_0^{t_1} d\tau. \tag{30}$$

We shall now compose the H function. Considering (6), we have

$$H = 2c_0 \int_t^{t_1} [\Psi(\tau) - Q_1(\tau)]\, k(\tau, t)\, u(t)\, d\tau + \\ + c_0 [\Psi(t) - Q_1(t)] + c_1 k(t_1, t)\, u(t) + c_2. \tag{31}$$

According to the transversality condition, $c_1 = 0$, $c_0 \leq 0$.

It is obvious from (31) that H has a maximum if

$$u(t) = \frac{1}{2}(N_1 + N_2) + \frac{1}{2}(N_2 - N_1) \operatorname{sign} \int_t^{t_1} [\Psi(\tau) - Q_1(\tau)]\, k(\tau, t)\, d\tau. \tag{32}$$

By substituting in this expression the $Q_1(\tau)$ value given by (30), we obtain an integral equation with respect to the function u(t) to be determined, the solution of which will define the extremum.

It is obvious from (32) that, in the case where $\Psi(\tau) \equiv Q_1(\tau)$, the first term of the H function becomes identically equal to zero, and that the determination of the optimum control function from the maximum principle becomes impossible. However, it is readily seen from (29) that the functional Q_0 to be minimized is identically equal to zero, i.e., it assumes the minimum possible value. According to the terminology given in [2], this case corresponds to the so-called "particular" trajectory section. The optimum control u = u(t) is, in this case, determined from the solution of the Volterra linear integral equation of the first kind:

$$\Psi(t) = \int_0^t k(t, \tau)\, u(\tau)\, d\tau. \tag{33}$$

The methods used for solving equations of the type given by (33) have been sufficiently treated in the literature [7].

Thus, on the basis of the above reasoning, we can find the optimum control inputs for a number of systems with distributed parameters.

APPENDIX I

<u>**Proof of Theorem 1.**</u> Let $u = u(t) \in \Omega$ $(t_0 \leq t \leq t_1)$ represent optimum control for which

$$Q_i = \int_{t_0}^{t_1} K_i(t, \tau\; u(\tau))\, d\tau \qquad (i = 1, 2, \ldots, n), \qquad Q(t_1) = Q_* \tag{34}$$

and the functional

$$Q_0 = \int_{t_0}^{t_1} F(\tau, Q(\tau), u(\tau))\, d\tau$$

assumes its minimum value.

Let us determine the control function $u^* = u^*(t)$, which is obtained by varying the control given by $u = u(t)$. For this, we shall select the instants of time $\tau_1, \tau_2, \ldots, \tau_l$ which satisfy the inequalities $t_0 < \tau_1 < \tau_2 < \ldots < \tau_l < t_1$, and which represent continuity points of the control function $u(t)$. We shall select arbitrary non-negative numbers $\delta t_1, \ldots, \delta t_l$, and arbitrary points $v_1, v_2, \ldots, v_l$, belonging to the control region Ω. Consider the half-intervals I_i given by

$$\tau_i - \varepsilon\delta t_i < t \leqslant \tau_i \quad (i = 1, 2, \ldots, l). \tag{35}$$

Assume that ϵ is so small that the half-intervals (35) never intersect each other, and that they are contained within the $t_0 \leq t \leq t_1$ section; then the $u^* = u^*(t)$ control will be determined in the following manner:

If t does not belong to any of the half-intervals $I_1, \ldots, I_l$, then

$$u^*(t) = u(t); \tag{36}$$

if $t \in I_i$ $(i = 1, 2, \ldots, l)$, then

$$u^*(t) = v_i. \tag{36'}$$

Then, the variated trajectory has the following form:

$$Q^*(\tau) = Q(\tau) + \varepsilon\Delta Q(\tau) + \ldots, \qquad Q^* = Q_0 + \varepsilon\Delta Q_0 + \ldots \tag{37}$$

Here, $\Delta Q(\tau)$ and ΔQ_0 are independent of ϵ, and they are given by

$$\Delta Q_i(\tau) = \sum_{j=1}^{l} [K_i(\tau, \tau_j, v_j) - K_i(\tau, \tau_j, u(\tau_j))]\,\delta t_j$$

$$\Delta Q_0 = \int_{t_0}^{t_1} \sum_{i=1}^{n'} \frac{\partial F}{\partial Q_i} \sum_{j=1} [K_i(\tau, \tau_j, v_j) - K_i(\tau, \tau_j, u(\tau_j))]\,\delta t_j d\tau + \tag{38}$$

$$+ \sum_{j=1}^{l} [F(\tau_j, Q(\tau_j), v_j) - F(\tau_j, Q(\tau_j), u(\tau_j))]\,\delta t_j.$$

Essentially, the $[\Delta Q_0, \Delta Q(t_1)]$ vector depends on the choice of the τ_i and v_i points, and the δt_i $(i = 1, 2, \ldots, l)$ numbers. (In the case where the same points τ_i correspond to different variations, control variations similar to those used in paragraph 9 of [1] can be introduced.) Moreover, it can be considered that the same number of τ_i and v_i points have been taken for different variations, so that the addition of new τ_i and v_i points for which $\delta t_i = 0$ would not change the control $u^*(t)$ that is being varied.

Let us now consider the multiplicity K of vectors $[\Delta Q_0, \Delta Q_0(t_1)]$ with the origin at the point $[Q_0, Q(t_1)]$ in space (Q_0, Q) which are obtained from the multiplicity of variations of the indicated type. We shall refer to the multiplicity set K by the term "cone of attainability." It can be readily seen that K is a convex cone. Actually, if the points s' and s'' belong to cone K, the point $s = \lambda_1 s' + \lambda_2 s''$, where $\lambda_1 \geq 0$, $\lambda_2 \geq 0$, also belongs to cone K, since the point s is obtained by variation which is determined by the joining of τ_i' and v_i' for s', and of τ_i'' and v_i'' for s'', while δt_i for the point s is equal to $\delta t_i = \lambda_1 \delta t_i'$ for τ_i', and to $\delta t_i = \lambda_2 \delta t_i''$ for τ_i''. Since Eqs. (38) are linear with respect to δt_i, we find that the point $s \in K$. Further, following a line of reasoning similar to that used in proving lemma 3 in [1], we find that the inside of cone K must not contain a straight line L with the origin at the apex of cone K, and which follows the negative direction of the Q_0 axis in the vector space (Q_0, Q). Hence, it follows that there are such numbers $c_0, c_1, \ldots, c_n$ for which the entire cone K lies in the half-space $\sum_{\alpha=0}^{n} c_\alpha Q_\alpha \leqslant 0$, and the straight line L lies in the half-space $\sum_{\alpha=0}^{n} c_\alpha Q_\alpha \geqslant 0$, i.e., the vector with the coordinates $c_0 = -1$, $c_1 = 0, \ldots, c_n = 0$ lies in the half-space $\sum_{\alpha=0}^{n} c_\alpha Q_\alpha \geqslant 0$. Consequently, $c_0 \leq 0$, whence we find that

$$\sum_{\alpha=0}^{n} c_\alpha \Delta Q_\alpha \leqslant 0, \; c_0 \leqslant 0. \tag{39}$$

Let us consider the $[\Delta Q_0, \Delta Q(t_1)]$ vector, which is obtained as a result of varying the control function $u(t)$ at the single point τ_1 $(l = 1)$ and $\delta t_1 = 1$. By taking (38) into account, we have

$$c_0\int_{t_0}^{t_1}\sum_{i=1}^{n}\frac{\partial F}{\partial Q_i}\left[K_i(\tau,\tau_1,v_1)-K_i(\tau,\tau_1,u(\tau_1))\right]d\tau+$$

$$+c_0[F(\tau_1,Q(\tau_1),v_1)-F(\tau_1,Q(\tau_1),u(\tau_1))]+\sum_{i=1}^{n}c_i\left[K_i(t_1,\tau_1,v_1)-K_i(t_1,\tau_1,u(\tau_1))\right]\leqslant 0. \tag{40}$$

Since this inequality is valid for any point $v_1 \in \Omega$, the function

$$H=c_0\int_{t_0}^{t_1}\sum_{i=1}^{n}\frac{\partial F}{\partial Q_i}K_i(\tau,t,u)\,d\tau+c_0F(t,Q(t),u)+\sum_{i=1}^{n}c_iK_i(t_1,t,u) \tag{41}$$

must have a maximum for $u \in \Omega$.

<u>Proof of Theorem 2.</u> Let us draw through the point $[Q_0, Q(t_1)]$ in space (Q_0, Q) a plane P that is parallel to space Q. We shall also draw a plane R with the dimension <u>n</u> (parallel to the Q_0 axis) through the plane S (with the $n-1$ dimension in space Q)–which is tangent to the multiplicity set M at the point $Q(t_1) \in M$– and the point $[Q_0, Q(t_1)]$ in space (Q_0, Q). Consider the half-plane T of the plane R that lies in that half of the space divided by the plane P where the negative "end" of the Q_0 axis is located. Due to the fact that the $Q = Q(t)$ $(t_0 \leq t \leq t_1)$ trajectory is the optimum trajectory, the attainability cone K must not contain any straight line L belonging to the half-plane T and having the origin at the point $[Q_0, Q(t_1)]$. In the opposite case, as at the beginning of the proof of lemma 10 from [1], we would arrive at a contradiction to the fact that $Q = Q(t)$ is not optimum, i.e., we could reach the multiplicity set M for a smaller value of Q_0. Cone K and the half-plane T represent convex cones with a common apex at the point $[Q_0, Q(t_1)]$, while the inside of cone K does not intersect with cone T. Therefore, there must exist a plane given by the equation $\sum_{\alpha=0}^{n} c_\alpha Q_\alpha = 0$ $(c_0 \leqslant 0)$, for which cone T lies in the half-space $\sum_{\alpha=0}^{n} c_\alpha Q_\alpha \geqslant 0$, and cone K lies in the half-space $\sum_{\alpha=0}^{n} c_\alpha Q_\alpha \leqslant 0$. For instance, the plane R, containing the half-plane T which, in turn, by definition, contains the plane S tangent to the multiplicity set M, is such a plane. This proves that the transversality conditions are satisfied.

APPENDIX II

By substituting (12) in (11), squaring, and by expanding into a sum of integrals, we have

$$I=\int_0^{x_1}q_a^2dx-2q_a\int_0^{x_1}\left(\int_0^{t_1}\varphi(x,t_1,\tau)u(\tau)d\tau\right)dx+\int_0^{x_1}\left(\int_0^{t_1}\varphi(x,t_1,\tau)u(\tau)d\tau\right)^2dx. \tag{42}$$

Let us reverse the order of integration in the second term of the sum, and let us denote $R(t,\tau)=\int_0^{x_1}\varphi(x,t,\tau)\,dx$ and $\gamma = 2q_a$. Let us represent the square of the integral with respect to time in the third term as a double integral, and also reverse the order of integration while denoting $T(t,\tau)=\int_0^{x_1}\varphi(x,t_1,t)\,\varphi(x,t_1,\tau)\,dx$. Then expression (42) will assume the following form:

$$I=x_1q_a^2-\gamma\int_0^{t_1}R(t_1,\tau)u(\tau)d\tau+\int_0^{t_1}\int_0^{t_1}T(t,\tau)u(t)u(\tau)\,dt\,d\tau. \tag{43}$$

It is obvious that (43) can be written thus:

$$I=x_1q_a^2+\int_0^{t_1}\left[\int_0^{t_1}T(t_1,\tau)u(t)\,dt-\gamma R(t_1,\tau)\right]u(\tau)\,d\tau. \tag{44}$$

By denoting $Q_1(t) = \int_0^{t_1} T(\tau, t)\, u(\tau)\, d\tau$ and neglecting the constant term $x_1 q_a^2$, we find that the minimization of functional I is equivalent to the minimization of the functional

$$Q_0 = \int_0^{t_1} [Q_1(\tau) - \gamma R(t_1, \tau)]\, u(\tau)\, d\tau \tag{45}$$

under the condition that

$$Q_1(t) = \int_0^{t_1} T(\tau, t)\, u(\tau), \qquad Q_2 = \int_0^{t_1} d\tau. \tag{46}$$

LITERATURE CITED

1. V. G. Boltyanskii, R. V. Gamkrelidze, and L. S. Pontryagin, "Theory of optimum processes. I," Izvest. Akad. Nauk SSSR, Math. Sec. 24, No. 1 (1960).
2. L. I. Rozonoér, "L. S. Pontryagin's maximum principle in the theory of optimal systems. I, II, and III," Avtomatika i Telemekhanika 20, Nos. 10, 11, 12 (1959).
3. A. G. Butkovskii and A. Ya. Lerner, "On the optimum control of systems with distributed parameters," Avtomatika i Telemekhanika 21, No. 6 (1960).
4. S. L. Sobolev, Mathematical Physics Equations [in Russian] (Gostekhizdat, 1954).
5. V. I. Smirnov, Calculus Course [in Russian] (Gostekhizdat, 1951) Vol. 4.
6. Z. S. Agranovich and A. Ya. Povzner, Application of Operational Methods to the Solution of Some Problems in Mathematical Physics [in Russian] (Khar'kov State University Press, 1954).
7. F. Trikomi, Integral Equations [Russian translation] (IL, Moscow, 1960).

CERTAIN PROBLEMS IN DESIGNING SYSTEMS WITH EXTREMAL SELF-ADJUSTING CORRECTIVE DEVICES

V. N. Varygin

Translated from Avtomatika i Telemekhanika, Vol. 22, No. 1, pp. 27-36, January, 1961

Original article submitted July 6, 1960

This article presents the derivation of equations of the extremum self-adjustment of corrective devices in linear systems for scanning oscillation frequencies that are comparable to the natural frequency of the basic control loop. These equations make it possible to investigate self-adjusting systems that are faster than those which can be investigated under quasi-steady-state conditions. The optimum relationships between the gain factors in self-adjustment loops are derived. Some of the results obtained in experimental investigation of such systems are given.

If elements whose parameters change at random (the system to be controlled, the regulator, etc.) are present in the automatic control loop, it often becomes necessary to solve the problem of the system dynamic characteristic stabilization by suitably varying the parameters of the correcting devices (CD). The principle of CD extremal adjustment can be used for this purpose [1].

The schematic diagram of such a system is shown in Fig. 1. The input signal θ is fed to the basic system W_b and to the reference system W_r, which is synthesized on the basis of the desired control performance criterion

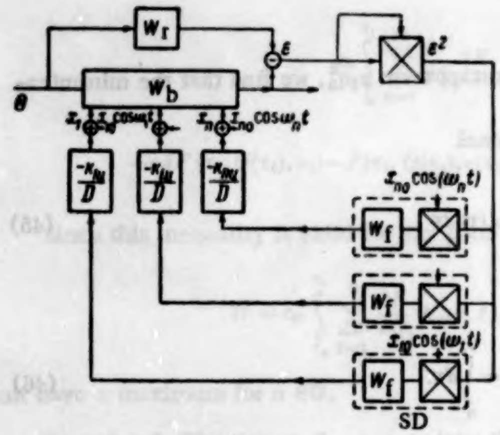


Fig. 1.

and the available information on the input signal. The difference ϵ between the reactions of the basic and the reference systems serves as the index of the deviation of the W_b system characteristics from the assigned characteristics. The quantity $\bar{\epsilon}^2$ can be conveniently used as the measure of this deviation.

The CD adjustment parameters $X_1, X_2, \ldots, X_n$ are subjected to random or regular scanning oscillations with a rather small amplitude. We shall consider only regular sinusoidal scanning oscillations* [7].

Synchronous detectors SD, which are connected to the error signal quadrator output, separate the components which are synchronous with the scanning oscillations, and which modify the average values of the adjustment parameters $X_1, X_2, \ldots, X_n$ through the integrating elements until the system arrives at the state corresponding to the minimum $\bar{\epsilon}^2$.

If the variations of the W_b system parameters are completely compensated by variations of the CD adjustment parameters, then, under steady-state conditions, the basic and the reference systems will be identical. Under transient conditions, the CD parameters will follow the variations of the basic system parameters with a certain dynamical error.

Equations of the Self-Adjustment Processes

The general first-approximation integral-differential equations describing the dynamics of the self-adjustment processes have been derived in [3]. These equations cannot be solved for the general case, and they are too complex for engineering purposes. Moreover, the above article considers the quasi-steady-state, self-adjustment conditions, which correspond to low scanning and self-adjustment rates in comparison with the processes taking place in the basic control loop. For this case, very simple equations, which make it possible to reduce investigations of the dynamics of such systems to the investigation of linear automatic control systems by ordinary methods, have been derived.

The scanning rate must be increased if the self-adjustment system operating speed is to be increased. This would make it possible to broaden the transmission band of filters at the output of the SD multiplying elements, which basically determine the speed of self-adjustment processes. The present paper is concerned with the above-described system for the case where the scanning oscillation frequencies are comparable to the natural frequency of the basic control loop.

For a small scanning oscillation amplitude and sufficiently slow changes in the values of the CD adjustment parameters, the system operator $W = W_r - W_b$ can be represented by the following fast-converging series (see Appendix):

$$W(D, x_1, x_2, \ldots, x_n, t) = W_0(D, X_1, X_2, \ldots, X_n, t) + \sum_{i=1}^{n} x_{i0}\cos\omega_i t W_i(D, X_1, X_2, \ldots, X_n, t) + \sum_{i=1}^{n} x_{i0}\omega_i \sin\omega_i t W_i'(D, X_1, X_2, \ldots, X_n, t) + P(D, X_1, X_2, \ldots, X_n, t), \tag{1}$$

where D is the differentiation symbol, ω_i is the scanning oscillation frequency with respect to the ith parameter, W_i and W_i' are the system operator components which take into account, in the first approximation, the presence of scanning oscillations with respect to the ith parameter (see Appendix), and

* The proposed investigation method can readily be extended to the case of periodic scanning oscillations of any form if they can be represented by a convergent Fourier series.

$$x_i = X_i + x_{i0}\cos\omega_i t, \quad x_{i0} \ll X_i,$$
$$W_0(D, X_1, X_2, \ldots, X_n, t) = W_9(D) - W_c(D, X_1, X_2, \ldots, X_n, t).$$

(The dependence on time is expressed by the fact that the $X_1, X_2, \ldots, X_n$ values are considered as "frozen" at the instant of time t.)

The $P(D, X_1, X_2, \ldots, X_n, t)$ operator includes the remaining terms of the series; since they include x_{i0} factors of higher power, i.e., terms which are small quantities of higher orders in comparison with W_i and W_i', they can be neglected.

The system weighting function can be written in the following form:

$$K(X_1, X_2, \ldots, X_n, t, \tau) = K_0(X_1, X_2, \ldots, X_n, t, \tau) + \sum_{i=1}^{n} x_{i0}\cos\omega_i t\, K_i(X_1, X_2, \ldots, X_n, t, \tau) + \sum_{i=1}^{n} x_{i0}\omega_i \sin\omega_i t\, K_i'(X_1, X_2, \ldots, X_n, t, \tau), \tag{2}$$

where $K_0(X_1, X_2, \ldots, X_n, t, \tau)$, $K_i(X_1, X_2, \ldots, X_n, t, \tau)$, and $K_i'(X_1, X_2, \ldots, X_n, t, \tau)$ correspond to the operators $W_0(D, X_1, X_2, \ldots, X_n, t)$, $W_i(D, X_1, X_2, \ldots, X_n, t)$, and $W_i'(D, X_1, X_2, \ldots, X_n, t)$.

Considering that the K_0, K_i, and K_i' weighting functions are analytical functions with respect to the adjustment parameters, they can be expanded into series which have the following forms:

$$\begin{aligned} K_0(X_1, X_2, \ldots, X_n, t, \tau) &= K_{0r}(t, \tau) + \sum_{j=1}^{n}\left(\frac{\partial K_0}{\partial X_j}\right)_r \Delta X_j + \ldots, \\ K_i(X_1, X_2, \ldots, X_n, t, \tau) &= K_{ir}(t, \tau) + \sum_{j=1}^{n}\left(\frac{\partial K_i}{\partial X_j}\right)_r \Delta X_j + \ldots, \\ K_i'(X_1, X_2, \ldots, X_n, t, \tau) &= K_{ir}'(t, \tau) + \sum_{j=1}^{n}\left(\frac{\partial K_i'}{\partial X_j}\right)_r \Delta X_j + \ldots, \end{aligned} \tag{3}$$

where $\Delta X_j = X_j - X_{jr}$, and X_{jr} is the adjustment parameter value corresponding to the steady-state conditions in the self-adjustment loops; the "r" index denotes the substitution of the

$$X_j = X_{jr} \quad (j = 1, 2, \ldots, n)$$

values in the expressions for the weighting functions and their derivatives.

According to design, the system must maintain the ΔX_j ($j = 1, 2, \ldots, n$) values close to zero; therefore, we shall hereafter consider only the first two terms of series (3).

By substituting (3) in (2), and neglecting terms which are small quantities of higher orders, we obtain

$$K(t, \tau) = K_{0r}(t, \tau) + \sum_{j=1}^{n}\left(\frac{\partial K_0}{\partial X_j}\right)_r \Delta X_j + \sum_{i=1}^{n} x_{i0}\cos\omega_i t\, K_{ir}(t, \tau) + \sum_{i=1}^{n} x_{i0}\omega_i \sin\omega_i t\, K_{ir}'(t, \tau). \tag{4}$$

If we assume that the variations of the average values of the adjustment parameters are inconsiderable within the range of large values of the basic loop weighting function $K(t)$ [6], we obtain

$$\varepsilon(t) = \int_{-\infty}^{\infty} K(t, \tau)\,\theta(\tau)\,d\tau, \tag{5}$$

$$\varepsilon^2(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K(t, \tau_1) K(t, \tau_2)\,\theta(\tau_1)\,\theta(\tau_2)\,d\tau_1\,d\tau_2. \tag{6}$$

Considering that the input signal is a random function of time, we shall write the $\theta(\tau_1)\theta(\tau_2)$ product in the following form [4]:

$$\theta(\tau_1)\,\theta(\tau_2) = R(\tau_1, \tau_2) + f(\tau_1, \tau_2), \tag{7}$$

where $R(\tau_1, \tau_2)$ is the input signal correlation function, and $f(\tau_1, \tau_2)$ is the centered random function.

By substituting (4) in (6), and by taking (7) into account, we obtain

$$\varepsilon^2(t) = 2\sum_{i=1}^{n} x_{i0}\cos\omega_i t\, S_i + 2\sum_{i=1}^{n} x_{i0}\omega_i \sin\omega_i t\, S'_i + R, \tag{8}$$

where

$$S_i = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left\{K_{0r}(t,\tau_1) + \sum_{j=1}^{n}\left[\frac{\partial K_0(t,\tau_1)}{\partial X_j}\right]_r\right\} K_{ir}(t,\tau_2)\,R(\tau_1,\tau_2)\,d\tau_1 d\tau_2, \tag{9}$$

$$S'_i = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left\{K_{0r}(t,\tau_1) + \sum_{j=1}^{n}\left[\frac{\partial K_0(t,\tau_1)}{\partial X_j}\right]\right\} K'_{ir}(t,\tau_2)\,R(\tau_1,\tau_2)\,d\tau_1 d\tau_2, \tag{10}$$

while R comprises the constant component, different combination frequencies, and the centered random functions.

The scanning oscillation frequencies must be chosen in such a manner that their combinations in the quadrator signal do not produce a constant component in multiplication by the SD reference voltages. At the output of the ith SD multiplying element, we have

$$u_i = x_{i0}\cos\omega_i t\,\varepsilon^2(t) = x_{i0}^2 S_i + Q_i, \tag{11}$$

where Q_i contains different, newly formed combination frequencies and centered random components.*

We shall assume that the linear portion of the self-adjustment circuits completely filters out the Q_i components. Actually, SD are resonance filters and, therefore, this condition can readily be satisfied by selecting the necessary operator W_f if the input signal θ spectrum does not have peaks near the scanning oscillation frequencies. Then, according to the diagram shown in Fig. 1,

$$\dot{X}_i = -k_i W_f(D)\,S_i \quad (i = 1, 2, \ldots, n), \tag{12}$$

where

$$k_i = k_{iu} x_{i0}^2.$$

Let us write S_i in the following form:

$$S_i = S_{0i} + \sum_{j=1}^{n}\Delta X_j S_{ij}, \tag{13}$$

where

$$S_{0i} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_{0r}(t,\tau_1)\,K_{ir}(t,\tau_2)\,R(\tau_1,\tau_2)\,d\tau_1\,d\tau_2, \tag{14}$$

$$S_{ij} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\frac{\partial K_0(t,\tau_1)}{\partial X_j}\right]_r K_{ir}(t,\tau_2)\,R(\tau_1,\tau_2)\,d\tau_1\,d\tau_2. \tag{15}$$

Then the system of equations (12) in expanded form can be written thus:

$$DX_1 = -k_1 W_f(D)\left[S_{01} + \sum_{j=1}^{n}\Delta X_j S_{1j}\right],$$

---

*Only one term of the ε^2 series, namely the $2x_{i0}\cos\omega_i t S_i$ term, produces in multiplication by $x_{i0}\cos\omega_i t$ a constant component that can be separated by the W_f filter.

$$DX_2 = -k_2 W_f\ (D)\Big[S_{02} + \sum_{j=1}^{n} \Delta X_j S_{2j}\Big],$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$DX_n = -k_n W_f\ (D)\Big[S_{0n} + \sum_{j=1}^{n} \Delta X_j S_{nj}\Big]. \tag{16}$$

Under steady-state conditions for the self-adjustment system, all ΔX_j (j = 1, 2, . . . , n) and all DX_i (i = = 1, 2, . . . , n) will be equal to zero. Then, the adjustment parameter values corresponding to the extremum can be determined from the system of equations

$$S_{0i} = 0 \quad (i = 1, 2, \ldots, n). \tag{17}$$

By taking into account (17) and the equation $DX_i = DX_{ir} + D\Delta X_i$, the characteristic equation of system (16) can be written in the following form:

$$\begin{vmatrix} \left[S_{11} + \frac{\lambda}{k_1 W_f\ (\lambda)}\right] & S_{12} & \cdots & S_{1n} \\ S_{21} & \left[S_{22} + \frac{\lambda}{k_2 W_f\ (\lambda)}\right] & \cdots & S_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ S_{n1} & S_{n2} & \cdots & \left[S_{nn} + \frac{\lambda}{k_n W_f\ (\lambda)}\right] \end{vmatrix} = 0. \tag{18}$$

The forms of Eqs. (17) and (18) are similar to those of the equations obtained in [3] for the quasi-steady-state conditions. The equations become identical if we assume that the scanning oscillation frequencies are infinitely small in comparison with the system natural frequency. In this case,

$$K_i(t, \tau) = \frac{\partial K_0(t, \tau)}{\partial X_i}, \quad S_{0i} = \frac{1}{2}\,\frac{\partial \overline{\varepsilon^2(t)}}{\partial X_i}, \quad S_{ij} = \frac{1}{2}\left(\frac{\partial^2 \overline{\varepsilon^2(t)}}{\partial X_i\, \partial X_j}\right)_r.$$

The characteristic equation (18) can be written in the following form:

$$\begin{vmatrix} k_1 S_{11} + I & k_1 S_{12} & \cdots & k_1 S_{1n} \\ k_2 S_{21} & k_2 S_{22} + I & \cdots & k_2 S_{2n} \\ k_n S_{n1} & k_n S_{n2} & \cdots & k_n S_{nn} + I \end{vmatrix} = 0, \tag{19}$$

where

$$I = \frac{\lambda}{W_f\ (\lambda)}. \tag{20}$$

If there is an extremum (a minimum), all I_i values are negative [5]. After determining the I_i (i = = 1, 2, . . . , n) values from the solution of the equation

$$I_i = \frac{\lambda}{W_f\ (\lambda)}, \tag{21}$$

the roots of the characteristic equation (18) can be found.

The solution of Eq. (21) can be reduced to the investigation of the stability of the hypothetical system shown in Fig. 2.

The characteristic equation of this system is identical to Eq. (21):

$$-I_i W_f\ (\lambda)\frac{1}{\lambda} + 1 = 0.$$

Thus, the stability condition for the system under consideration can be defined (for $I_i < 0$ for all i) as the stability condition for an isolated channel (Fig. 2) for $I_i = I_{max}$. Actually, if the real parts of the roots of Eq. (21) are negative for $I_i = I_{max}$, they will be certainly negative for all $I_i < I_{max}$.

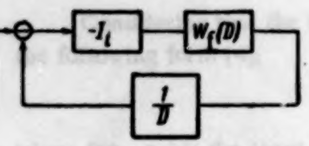


Fig. 2.

The control time will be defined by

$$t_{reg} \approx 3 \frac{1}{\lambda_{min}}, \tag{22}$$

where λ_{min} is the root with the smallest absolute value which is found from Eq. (21) for $I_i = I_{min}$ (if the roots are real; if the roots are complex, the smallest value of the root real part must be used instead of λ_{min}).

If the difference between I_{max} and I_{min} is large, the processes will be delayed, since the W_f filter gain, which is determined from the stability condition for Eq. (21) for $I_i = I_{max}$, will yield insufficiently large root values in substituting $I_i = I_{min}$. By determining the optimum relationship between $k_1, k_2, \ldots, k_n$, the ratio I_{max}/I_{min} can be reduced to the minimum.*

Let us find the optimum relationships between the k_1 and k_2 coefficients from the condition for the I_{max}/I_{min} minimum for a system with two adjustable parameters. The characteristic equation for this case will have the following form:

$$\begin{vmatrix} k_1S_{11} + I & k_1S_{12} \\ k_2S_{21} & k_2S_{22} + I \end{vmatrix} = 0.$$

Hence,

$$I_{1,2} = -\frac{k_1S_{11} + k_2S_{22}}{2} \pm \frac{1}{2}\sqrt{(k_1S_{11} + k_2S_{22})^2 + 4k_1k_2S_{12}S_{21} - 4k_1k_2S_{11}S_{22}}.$$

It can be shown that the minimum of I_1/I_2 with respect to k_1/k_2 will be obtained for $k_1/k_2 = S_{22}/S_{11}$, i.e., for $k_1S_{11} = k_2S_{22}$.

This is apparently the optimum relation also for a large number of adjustable parameters, i.e., for any n,

$$k_1S_{11} = k_2S_{22} = \ldots = k_nS_{nn} \tag{23}$$

must hold.

The performed experiments indicate that the use of condition (23) leads to a tenfold, and sometimes hundredfold, increase in the self-adjustment process speed in comparison with the case where the channel gains are equal.

If the I_{max} and I_{min} values cannot be brought closer together even for the optimum k_i value, the method of coupled control is recommended [1].

Introduction of Scanning Oscillations into the Reference System

In a number of cases, the introduction of scanning oscillations into the basic system either is inadmissible, or it entails considerable technical difficulties. Then, the application of scanning oscillations to the corresponding reference system parameter can be recommended. Such scanning has a number of important advantages:

1. The scanning oscillation amplitude can be made much larger than the amplitude in the previous scanning variant, which would improve the self-adjusting system stability and noise-stability.

2. No distortions due to scanning oscillations would be present in the main system during the signal transmission through the basic loop.

3. By arranging the tracking of parameter changes in the basic system by the parameters of the reference system (the basic system model), and by introducing scanning oscillations into the reference model, the solutions of a number of purely extremal problems, which cannot be solved by ordinary methods, can be found.

A disadvantage of the method of introducing scanning oscillations into the reference system is the necessity of designing reference systems with variable parameters.

*In the limiting case, it can be reduced to unity.

X_{in}, v	t_{reg}, sec Calculated from (22)	Obtained experimentally
3	45	45
4	23	22
5	13	12

In conclusion, it should be noted that, on the basis of the experience gained in practical work, the equations derived under the assumption that the self-adjustment processes are slow in comparison with the basic loop operating speed are also valid with sufficient accuracy in the case where the basic loop operating speed is comparable to that of the self-adjustment loop.

The author hereby expresses his gratitude to G. Yu. Polkanov for his great help in performing the experiments.

Example. Let us consider the system for the transmission band stabilization in a very simple servosystem whose equation is given by

$$\dot{y} + ky = kx. \tag{24}$$

The gain is subjected to periodic scanning oscillations about a certain average value $\bar{k} = k_a$, i.e.,

$$k = k_a + k_0 \cos \omega_0 t \tag{25}$$

for $k_0 \ll k_a$.

With sufficient accuracy, Eq. (24) can be written as $\dot{y}(T_a - T_0 \cos \omega_0 t) + y = x$, where

$$T_a = \frac{1}{k_a}, \quad T_0 = \frac{k_0}{k_a^2}. \tag{26}$$

In this case, we shall use the following inertial element with the required time constant T_r as the reference system:

$$W_r(D) = \frac{1}{T_r D + 1}. \tag{27}$$

By using (1), we shall find the basic system operator

$$W_a(D) \approx \frac{1}{T_a D + 1} + \frac{T_0 \cos \omega_0 t\, D}{(T_a D + 1)^2 + T_a^2 \omega_0^2} + \frac{\omega_0 T_a T_0 \sin \omega_0 t\, D}{[(T_a D + 1)^2 + T_a^2 \omega_0^2](T_a D + 1)} \tag{28}$$

and

$$W(D) = W_r(D) - W_a(D) \approx \frac{\Delta T D}{(T_r D + 1)^2} -$$

$$-T_0 \cos \omega_0 T \frac{D}{(T_r D + 1)^2 + T_r^2 \omega_0^2} - \sin \omega_0 t \frac{D T_0 T_a \omega_0}{[(T_r D + 1)^2 + T_r^2 \omega_0^2](T_r D + 1)}. \tag{29}$$

Considering (4), we can write

$$K_{0r}(t) = \frac{\Delta T}{T_r^3}(T_r - t)\, e^{-\frac{t}{T_r}},$$

$$K_{1r}(t) = \frac{T_0}{T_r^3} \sqrt{T_r^2 \omega_0^2 + 1}\, e^{-\frac{t}{T_r}} \cos(\omega_0 t + \varphi), \tag{30}$$

where

$$\varphi = \tan^{-1} \frac{1}{T_r \omega_0}.$$

Hence, S_{01} and S_{11} can be determined by using Eqs. (14) and (15). For an input signal $\theta = X_{in} \cos \omega t$, we obtain

$$S_{01} = \frac{[T_a - T_r][(1 - T_r^2 \omega_0^2)(1 - T_r^2 \omega^2 + T_r^2 \omega_0^2) + 4T_r^2 \omega^2]\, \omega^2}{2[4T_r^2 \omega^2 + (1 - T_r^2 \omega^2 + T_r^2 \omega_0^2)](T_r^2 \omega^2 + 1)} X_{in}^2, \tag{31}$$

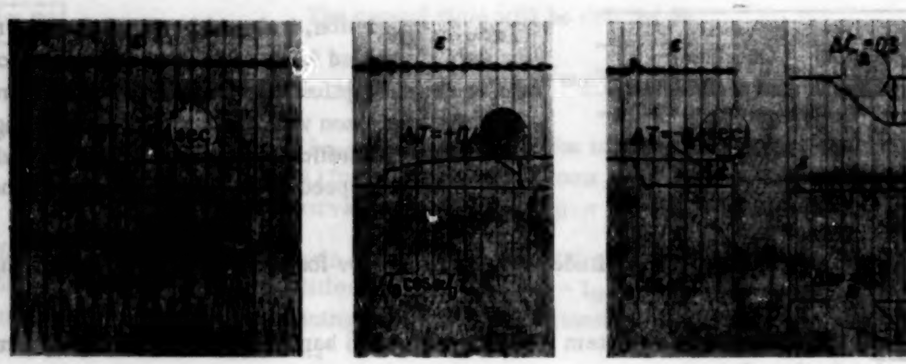


Fig. 3.

$$S_{11} = \frac{[(1 - T_r^2\omega^2)(1 - T_r^2\omega^2 + T_r^2\omega_0^2) + 4T_r^2\omega^2]\,\omega^2}{2\,[4T_r^2\omega^2 + (1 - T_r^2\omega^2 + T_r^2\omega_0^2)]\,(T_r^2\omega^2 + 1)}\,X_{in}^2. \tag{32}$$

The value $T_{sts} = T_r$, found from the solution of the equation $S_{01} = 0$, has been substituted in expression (32).

In this case, the self-adjustment process characteristic equation has the following form:

$$S_{11} - \frac{\lambda}{k_1 W_f\ (\lambda)} = 0. \tag{33}$$

By substituting the expression for W_f in Eq. (33), we can find all roots of the self-adjustment process characteristic equation.

The oscillograms in Figs. 3a,b, and c illustrate the above example for the following data:

$$T_r = 1\ \text{sec};\ T_0 = 0.1\ \text{sec};$$

$$\omega_0 = 1.26\ \text{sec}^{-1};\ \omega = 6.28\ \text{sec}^{-1};$$

$$W_f\,(D) = \frac{1}{1.5D + 1};$$

$X_{in} = 3$ v (for Fig. 3a); $X_{in} = 4$ v (for Fig. 3b); $X_{in} = 5$ v (for Fig. 3c).

The regulating time t_{reg} for the initial deviation from the extremum, calculated by using formula (22), and obtained experimentally, is given in the table.

The possibility of adjusting two parameters is illustrated in the oscillogram given in Fig. 3d, which shows the transient processes in regulating the damping decrement and the natural frequency of an oscillatory element with the transfer function

$$W_a = \frac{\omega_a^2}{D^2 + 2\omega\,\xi D + \omega_a^2}$$

for

$$W_r = \frac{\omega_{ar}^2}{D^2 + 2\xi_r\omega_{ar}D + \omega_{ar}^2}$$

and

$$\xi_r = 0.7,\quad \omega_{ar} = 1\ \text{sec}^{-1}$$

in the case of initial deviation of ω and ξ_a from ω_{ar} and ξ_r.

APPENDIX

Determination of the Transfer Locus of a System with Extremal Self-Adjustment of the Corrective Devices

The system equation with variable coefficients has the following form:

$$[a_n(t)D^n + \ldots + a_1(t)D + a_0(t)]\,y(t) = [b_m(t)D^m + \ldots + b_1(t)D + b_0(t)]\,x(t) \tag{34}$$

or

$$L(D,\,t)\,y(t) = K(D,\,t)\,x(t), \tag{35}$$

where x(t) and y(t) are the input and output signals, respectively; a(t) and b(t) are certain known time functions, and D is the differentiation symbol.

In [6], frequency methods, which are widely used in investigating systems with constant parameters, have been applied to systems with variable parameters. The $H(j\omega, t)$ function, which is here used as the frequency response analog, is found by solving the following equation:

$$[\alpha_n(t)D^n + \ldots + \alpha_1(t)D + \alpha_0(t)]\,H(j\omega,\,t) = K(j\omega,\,t), \tag{36}$$

where $\alpha_\mu(t) = \frac{1}{\mu!}\cdot\frac{\partial^\mu L(j\omega,\,t)}{\partial (j\omega)^\mu}$, and $K(j\omega, t)$, $L(j\omega, t)$ are obtained by substituting $j\omega$ for D in the expressions for K(D, t) and L(D, t).

By using the method proposed in [6], the solution of Eq. (36) is obtained as a series with the following form:

$$H(j\omega,\,t) = H_1(j\omega,\,t) + H_2(j\omega,\,t) + \ldots, \tag{37}$$

which is a fast-converging series if the coefficients of Eq. (36) oscillate about their average values, which actually takes place in the system considered in the main portion of the present article.

For determining the terms of series (37), Eq. (36) should be rewritten thus:

$$[\bar{\alpha}_n D^n + \ldots + \bar{\alpha}_1 D + \bar{\alpha}_0]\,H = K + P(H), \tag{38}$$

where $\bar{\alpha}_\mu$ is the average value of $\alpha_\mu(t)$ [or a constant approximating $\alpha_\mu(t)$], and P(H) is the totality of variable terms in the equation, i.e.,

$$P(H) = H\,\{[\bar{\alpha}_n - \alpha_n(t)]\,D^n + \ldots + [\bar{\alpha}_1 - \alpha_1(t)]\,D + [\bar{\alpha}_0 - \alpha_0(t)]\}. \tag{39}$$

The steady-state solution of the following equation with constant coefficients will represent the first approximation of $H(j\omega, t)$:

$$[\bar{\alpha}_n D^n + \ldots + \bar{\alpha}_1 D + \bar{\alpha}_0]\,H_1 = K. \tag{40}$$

The $H_2, H_3, \ldots$ values are obtained from the steady-state solutions of the equations

$$[\bar{\alpha}_n D^n + \ldots + \bar{\alpha}_1 D + \bar{\alpha}_0]\,H_\mu = P(H_{\mu-1}) \tag{41}$$

In our case, if we consider the accepted assumption that the average values of the adjustable parameters change little in the region of large values of the basic control loop weighting function, and that the scanning oscillation amplitude is small, it is sufficient to determine only the first two terms of series (37).

As an illustration of this, consider the above example, where

$$K(j\omega,\,t) = 1,\; L(j\omega,\,t) = j\omega\,(T_a - T_0\cos\omega_0 t) + 1.$$

For this case, Eq. (38) will assume the following form:

$$\left(\frac{T_a}{T_a j\omega+1}D+1\right)H=\frac{1}{j\omega T_a+1}+\cos\,\omega_0 t\,\frac{T_0(1+D)}{j\omega T_a+1}H.$$

Then, according to (40) and (41),

$$H_1=\frac{1}{T_a j\omega+1},$$

$$H_2=T\cos\omega_0 t\,\frac{j\omega}{(T_a j\omega+1)^2+T_a^2\omega_0^2}+T_0\omega_0\sin\omega_0 t+\frac{j\omega T_a}{[(T_a j\omega+1)^2+T_a^2\omega_0^2](T_a j\omega+1)}.$$

The next term of the series will be found by solving the equation

$$(T_a D+1)H_3=\cos\omega_0 t\,\frac{T_0}{j\omega T_a+1}H_2$$

and, since it will contain the factor T_0^2, it can be neglected.

Every adjustment parameter in Eq. (1) corresponds to a pair of $x_{i0}\cos\omega_i t W_i(D, X_1, X_2, \ldots, X_n, t)$ and $x_{i0}\sin\omega_0 t W_i^*(D, X_1, X_2, \ldots, X_n, t)$ terms, the origin of which is clear from the above example, where

$$x_{i0}=T_0,\; W_i=\frac{j\omega}{(T_a j\omega+1)^2+T_a^2\omega_0^2}\;;\; W_i^*=\frac{j\omega T_a}{[(T_a j\omega+1)^2+T_a^2\omega_0^2](T_a j\omega+1)}.$$

The substitution of D for $j\omega$ in (1) is possible, since W_i and W_i^* represent the transfer functions of a system with constant parameters.

LITERATURE CITED

1. A. A. Krasovskii, Self-Adjusting Systems in Automatic Control [in Russian] (N. E. Zhukovskii VVIA Press, 1959).
2. A. A. Krasovskii, "Dynamics of continuous extremal automatic control systems which are based on the gradient method," Izvest. Akad. Nauk SSSR, Otd. Tekh. N. No. 3 (1959).
3. A. A. Krasovskii, "Dynamics of continuous automatic control systems with extremal self-adjustment of the corrective devices," Report to the First International Congress on Automatic Control (1960).
4. J. H. Laning and R. G. Battin, Stochastic Processes in Automatic Control Problems [Russian translation] (IL, 1958).
5. N. G. Chetaev, Motion Stability [in Russian] (Gostekhizdat, 1949).
6. L. Zadeh, "Frequency analysis of variable networks," Proc. IRE 38, No. 3 (1950).
7. S. Draper and J. T. Li, Principles of Optimizing Control Systems and an Application to Internal Combustion Engines (ASME Publications, 1951).

All abbreviations of periodicals in the above bibliography are letter-by-letter transliterations of the abbreviations as given in the original Russian journal. *Some or all of this periodical literature may well be available in English translation.* A complete list of the cover-to-cover English translations appears at the back of this issue.

DETERMINATION OF THE OPTIMUM VARIATION MODE
OF THE USEFUL SIGNAL AND NOISE CARRIER FREQUENCIES
IN DETECTION PROBLEMS BASED ON THE THEORY OF GAMES

M. Yu. Gadzhiev
Translated from Avtomatika i Telemekhanika, Vol. 22, No. 1,
pp. 37-46, January, 1961
Original article submitted May 16, 1960

The conflicting situation arising when the useful signal is received against the noise background, in the case where the useful signal carrier frequencies and the noise can change within the assigned frequency range, is considered here. The possible payoff function for the game model of the situation is determined, and the physical meaning of the payoff function is defined. The payoff function is approximated by expressions which make it possible to find the optimal strategy and the game value by means of the existing methods. Concrete examples are given.

The natural course to be followed in noise interference control is that of attempting to tune it out by changing the useful signal carrier frequency. If, after the useful signal frequency has been changed, the noise carrier frequency is changed in such a manner that the most effective interference by noise is secured, a conflicting situation arises, where noise and the useful signal have opposite "interests."

Under certain simplifying assumptions, this situation can be represented in terms of the theory of games [1,2] by assigning the role of one (winning) player to noise, and the role of the other (losing) player to the useful signal. In the game corresponding to the conflicting situation under consideration, either of the players has only one personal move: the independent and simultaneous selection of his carrier frequency value from the closed interval $[W_1, W_2]$.

Let us introduce the basic concepts and terms of the theory of games as applied to the problem under consideration: f_1 is the pure strategy of the first player (noise), which consists in the selection of the carrier frequency f_1; f_2 is the pure strategy of the other player (useful signal), which consists in the selection of the frequency f_2; $\Phi(f_1, f_2)$ is the payoff function; $\xi(f_1)$ is the mixed strategy of the first player; $\eta(f_2)$ is the mixed strategy of the second player; $\xi_0(f_1)$ is the optimal mixed strategy of the first player; $\eta_0(f_2)$ is the optimal mixed strategy of the second player; ν is the game value.

In operation, the useful signal receiver is approximated by a linear weighting function. In the general case, the noise state is assumed to be unsteady.

Selection of the Payoff Function

Let the useful signal $\varphi(t)$ become mixed with noise $X_0(t)$, thereby forming the signal

$$Z(t) = X_0(t) + \varphi(t), \tag{1}$$

which arrives at the receiver.

Let us denote

$$X_0(t) = X(t) \exp i2\pi f_1 t = N(t) \exp i2\pi [f_1 t + \alpha], \tag{2}$$

where N(t) and α are the envelope and the phase of high-frequency oscillation (2), respectively;

$$R_0(t_1, t_2) = R(t_1, t_2) \exp i\, 2\pi f_1 (t_1 - t_2) \tag{3}$$

is the noise complex correlation function;

$$g_0(t) = g(t) \exp i\,2\pi f_2 t = h(t) \exp i\,2\pi [f_2 t + \beta] \tag{4}$$

is the weighting function of the useful signal linear receiver; h(t) and β are the envelope and the phase of the weighting function (4), respectively.

We shall use the following expression for the payoff function:

$$\Phi(f_1, f_2) = \overline{|v(f_1, f_2)|^2} = \overline{\left[\int_{s-T}^{s} X_0(t)\, g_0^*(t)\, dt\right]^2} =$$

$$= \int_{s-T}^{s}\int_{s-T}^{s} R_0(t_1, t_2)\, g_0^*(t_1)\, g_0(t_2)\, dt_1\, dt_2, \tag{5}$$

where (s − T, s) is the interval of integration by the receiver of signal (1); the asterisk denotes the complex conjugate of the corresponding quantity, and the bar on top denotes statistical averaging.

By taking into account (2) and (4), the payoff function (5) can be written in the following form*:

$$\Phi(f_1, f_2) = \int \int_{s-T}^{s} R(t_1, t_2)\, h(t_1)\, h(t_2) \exp[i2\pi\Delta f(t_1 - t_2)]\, dt_1\, dt_2 =$$

$$= \int_{s-T}^{s}\int_{s-T}^{s} R(t_1, t_2)\, h(t_1)\, h(t_2) \cos[2\pi\Delta f(t_1 - t_2)]\, dt_1\, dt_2 = \Phi(f_1 - f_2), \tag{6}$$

where $\Delta f = f_1 - f_2$.

Let us determine the possible physical meaning of the payoff function (5). We shall demonstrate that the payoff function (5) is a generalization of the case of unsteady-state noise, where the possibility of changes in the carrier frequencies to be effected by the players, the so-called "noise function," is taken into account. This function has been investigated by Woodward [3] for white noise and for the coincidence of the carrier frequencies.

Let us consider the problem [3] of determining the "essential" parameter U of the $\varphi(t, U)$ signal. It is known [4] that, for extracting the useful information on U that is contained in Z, the optimal receiver must perform the following operation:

$$\eta(U, U') = \mathrm{Re}\int_{s-T}^{s} Z(t, U)\, g_n^*(t, U')\, dt =$$

$$= \mathrm{Re}\int_{s-T}^{s} \varphi(t, U)\, g_n^*(t, U')\, dt + \mathrm{Re}\int_{s-T}^{s} X_0(t)\, g_n^*(t, U')\, dt, \tag{7}$$

where U' is the assumed value of the "essential" parameter U. The optimum weighting function $g_n(t)$ will be determined by solving the equation

$$\int_{s-T}^{s} R_0(t_1, t_2)\, g_n(t_2, U')\, dt_2 = \varphi(t_1, U'). \tag{8}$$

---

* If the averaging is performed with respect to the phases of high-frequency oscillations (2) and (4), it can readily be shown that

$$\overline{[\mathrm{Re}\, v(f_1 - f_2)]^2} = \overline{[\mathrm{Im}\, v(f_1 - f_2)]^2} = \frac{1}{2}\Phi(f_1 - f_2),$$

where Re and Im are the real and the imaginary parts of the complex quantity.

while the solution is such that

$$g_n(t) \begin{cases} \not\equiv 0 & \text{for } s-T \leqslant t \leqslant s, \\ \equiv 0 & \text{for } s-T > t > s. \end{cases} \tag{9}$$

Let us consider the terms in Eq. (7), where $\operatorname{Re} \int_{s-T}^{s} \varphi(t, U)\, g_n^*(t, U')\, dt = f(U, U')$ is the so-called signal function, and $\operatorname{Re} \int_{s-T}^{s} X_0(t)\, g_n^*(t, U')\, dt = q(U')$ is the noise function.

The U parameter is evaluated with respect to the signal function maximum [3], which is attained for U=U':

$$\max f(U, U') = m(U) = \int_{s-T}^{s} \varphi(t, U)\, g_n^*(t, U)\, dt. \tag{10}$$

The error of the parameter U determination can be estimated with respect to the value of the noise function mean square:

$$\overline{|q|^2} = n(U') = \int_{s-T}^{s} \int_{s-T}^{s} R_0(t_1, t_2)\, g_n^*(t_1, U')\, g_n(t_2, U')\, dt_1\, dt_2 =$$

$$= 2\,\overline{[\operatorname{Re} q]^2} = 2\,\overline{[\operatorname{Im} q]^2}. \tag{11}$$

Expression (11) coincides with the payoff function (5).*

Let us define more accurately the payoff function (6) for the particular case of steady-state noise by changing the integration limits to $(-\infty, \infty)$ in (5) in correspondence with formula (9):

$$\Phi(\Delta f) = \int_{-\infty}^{\infty} R(\alpha)\, H(\alpha)\, [\cos 2\pi\Delta f\alpha]\, d\alpha. \tag{12}$$

Here,

$$H(\alpha) = \int_{-\infty}^{\infty} h(t)\, h(t+\alpha)\, dt = \int_{-\infty}^{\infty} K(f) \exp(i2\pi f\alpha)\, df =$$

$$= \int_{-\infty}^{\infty} |P(f)|^2 \exp(i2\pi f\alpha)\, df. \tag{13}$$

If $S(f) = \int_{-\infty}^{\infty} R(\alpha) \exp(-i2\pi f\alpha)\, d\alpha$ is the noise spectral density, then, with the new notation, Eq. (12) will assume the following form:

$$\Phi(\Delta f) = \int_{-\infty}^{\infty} K(f)\, \frac{S(f+\Delta f) + S(f-\Delta f)}{2}\, df. \tag{14}$$

After the substitution of variables, Eq. (14) can be written as

$$\Phi(\Delta f) = \int_{-\infty}^{\infty} \frac{K(f+\Delta f) + K(f-\Delta f)}{2}\, S(f)\, df. \tag{15}$$

---

*By using (8) and (9), it can be shown that the signal-function maximum is equal to the mean square of the noise function, i.e.,

$$n(U') = m(U).$$

Moreover, the noise function mean square is equal to the generalized signal-to-noise ratio [5] in using the optimum filter, which is determined from Eq. (8).

It can be demonstrated that the payoff function (14) or (15) represents the double output of a linear receiver with the frequency response $K(|f| - f_2)$ if narrow-band noise with the spectral density $S(|f| - f_1)$ is fed to its input.

Two particular cases of Eqs. (14) and (15) are given below:

The case of broad-band noise

$$\Phi(\Delta f) = S(\Delta f) \int_{-\infty}^{\infty} K(f)\, df. \tag{16}$$

With an accuracy to the constant factor, the payoff function (16) reproduces the noise power spectral density.

The case of narrow-band noise

$$\Phi(\Delta f) = K(\Delta f) \int_{-\infty}^{\infty} S(f)\, df. \tag{17}$$

Let us consider certain properties of the chosen payoff function.

1. It is natural to assume that $R(\alpha)H(\alpha)$ in (12) tends to zero when the argument α tends to $+\infty$ or $-\infty$. Due to the fact that $R(\alpha)H(\alpha)$ is an even function, $R(\alpha)H(\alpha)$ represents the Fourier transformation for the payoff function (12). Therefore, for $\Delta f \to \pm\infty$, the payoff function (12) tends to zero (perhaps while oscillating in the process). It is assumed here that the payoff function (6) also exhibits such an asymptotic behavior for large Δf values.

2. By definition, the payoff function is nonnegative and is an even function with respect to Δf, i.e.,

$$\Phi(\Delta f) = \Phi(-\Delta f) \geqslant 0.$$

3. It follows from the physical meaning of the payoff function that the first player (noise) must choose such carrier frequencies as to increase the values of the payoff function (6) or (12), and that the other player must strive to reduce these values by a suitable choice of carrier frequencies.

4. By differentiation with respect to f_1 and f_2 in the $f_1 \leq f_2$ and $f_1 \geq f_2$ regions, it can be established that, if payoff function $\Phi(f_1 - f_2)$ is nondecreasing with respect to f_1 in the $f_1 \leq f_2$ region, it is nonincreasing with respect to f_2 in the same region; in the $f_1 \geq f_2$ region, the $\Phi(f_1 - f_2)$ function is nonincreasing with respect to f_1 and nondecreasing with respect to f_2.

Therefore, by taking into account what has been said in paragraph 3, we arrive at the following method for manipulating the carrier frequencies, which is actually used in practice.

While the carrier frequencies are being varied by the players, the noise must strive to reduce the difference between its own carrier frequency and the useful signal-carrier frequency and, conversely, this difference must be increased for a better useful signal separation.

Game Solution Methods

The relationship between the payoff function analytical properties and the players' optimal strategies is given in [7, 8].

Thus, if the payoff function consists of a "bell-shaped" function, which is, generally speaking, characterized as positive, analytical, and monotonically decreasing to zero if its argument increases at a rate not lower than that of the exponent, the players' strategies are given by step functions [7].

The Gauss curve

$$\Phi(f_1 - f_2) = \exp[-\gamma(f_1 - f_2)^2]$$

can serve as an example of such a payoff function.

Optimal strategies have also been found for payoff functions of the Green function type [8]. Such functions are characterized by the fact that they are nonanalytical (it is usually assumed that there is a jump of the first derivative on the main diagonal of the square on which the payoff function is given, i.e., for $f_1 = f_2$). In this case, the players' optimal strategies are formed from the continuous component that is different from zero throughout the entire $[W_1, W_2]$ range, and the jumps at the W_1 and the W_2 boundaries.

We shall now consider the application of general game solution methods [1,2] to the payoff functions (6) and (12), without imposing on the payoff function the special restrictions which were necessary in [7,8].

1. The theory of games provides the proof [1] that the solutions of the equations

$$\int_{W_1}^{W_2} \Phi(f_1 - f_2)\,\xi'(f_1)\,df_1 = \nu', \tag{18}$$

$$\int_{W_1}^{W_2} \Phi(f_1 - f_2)\,\eta'(f_2)\,df_2 = \nu' \tag{19}$$

with the additional normalizing conditions

$$\int_{W_1}^{W_2} \xi'(f_1)\,df_1 = \int_{W_1}^{W_2} \eta'(f_2)\,df_2 = 1 \tag{20}$$

determine the optimal mixed strategies and the game value, i.e.,

$$\xi'(f_1) = \xi_0(f_1), \qquad \eta'(f_2) = \eta_0(f_2), \qquad \nu' = \nu$$

in the case where

$$\xi'(f_1) > 0, \qquad \eta'(f_2) > 0 \tag{21}$$

for all f_1 and f_2 contained in $[W_1, W_2]$.

Equations (18) and (19) are Wiener-Hopf equations with finite integration limits and with a kernel which is an even function of the argument difference. Therefore, if the payoff function for the problem under consideration is approximated by the sum

$$\Phi(f_1 - f_2) = \sum_{k=1}^{W} \varphi_k \exp(-\gamma_k |f_1 - f_2|), \tag{22}$$

where φ_k are constants and $\operatorname{Re}\gamma_k > 0$ ($k = 1, \ldots, N$), then, as is known, the solution of the Wiener-Hopf equation is obtained in closed form [6]. After solving Eqs. (18) and (19), we must check whether condition (21) is satisfied for each concrete payoff function (22).

Assume that the Fourier transformation of expression (22) is given by

$$G(\tau^2) = \sum_{k=1}^{N} \frac{2\gamma_k \varphi_k}{\tau^2 + \gamma_k^2} = \theta_0 \frac{\prod_{s=1}^{N-1}(\tau^2 + \alpha_s^2)}{\prod_{k=1}^{N}(\tau^2 + \gamma_k^2)} =$$

$$= \theta_0 \frac{\prod_{s=1}^{N-1}(\tau^2 + \alpha_s^2)}{(\tau^2 + \gamma_r^2)\prod_{s=1}^{N-1}(\tau^2 + \gamma_s^2)} = \frac{\sum_{s=1}^{N} \theta_{2s}\tau^{2(N-s)}}{\tau^{2N} + \sum_{k=1}^{N} \lambda_{2k}\tau^{2(N-k)}}. \tag{23}$$

The solution of Eq. (18) [since Eqs. (18) and (19) are identical, the solutions will hereafter be written only for (18)] will then be written in the following manner [6]:

$$\xi(f_1) = A\delta(f_1 - W_1) + B\delta(f_1 - W_2) + \\ + \sum_{j=1}^{N-1} \{C_j \exp[-\gamma_j(f_1 - W_1)] + D_j \exp[-\gamma_j(W_2 - f_1)]\} + \zeta_p(f_1), \tag{24}$$

where δ is the Dirac delta-function, and $\zeta_p(f_1)$ is the solution of the nonhomogeneous equation

$$G\left[-\left(\frac{d}{df}\right)^2\right]\zeta_p(f_1) = \nu. \tag{25}$$

For the case under consideration, this equation assumes the following form:

$$\zeta_p = \frac{\lambda_{2N}}{\theta_{2N}}. \tag{26}$$

By using, for instance, the method proposed in [6], it can be shown that the constants in (24) satisfy the following relations:

$$A = B = a\nu, \quad C_j = D_j = c_j\nu \quad (j = 1, \ldots, N-1), \tag{27}$$

where $\underline{a}$ and c_j are determined by the poles and roots of $G(\tau^2)$.

Thus, the final equations for the optimal mixed strategies of the players will be given by

$$\xi(f_1) = \nu\Big\{a[\delta(f_1 - W_1) + \delta(f_1 - W_2)] + \\ + \sum_{j=1}^{N-1} c_j[\exp[-\gamma_j(f_1 - W_1)] + \exp[-\gamma_j(W_2 - f_1)]] + \frac{\lambda_{2N}}{\theta_{2N}}\Big\}, \tag{28}$$

$$\eta(f_2) = \nu\Big\{a[\delta(f_2 - W_1) + \delta(f_2 - W_2)] + \\ + \sum_{j=1}^{N-1} c_j[\exp[-\gamma_j(f_2 - W_1)] + \exp[-\gamma_j(W_2 - f_2)]] + \frac{\lambda_{2N}}{\theta_{2N}}\Big\}. \tag{29}$$

The game value, determined from the normalizing condition (20), is equal to

$$\nu = \frac{1}{2a_1 + \sum_{j=1}^{N-1} 2\frac{b_j}{\gamma_j}\{1 - \exp[-\gamma_j(W_2 - W_1)]\} + \frac{\lambda_{2N}}{\theta_{2N}}(W_2 - W_1)}. \tag{30}$$

For instance,

$$\Phi(f_1 - f_2) = \varphi\exp(-\gamma|f_1 - f_2|), \tag{31}$$

the optimal mixed strategies are

$$\xi_0(f_1) = \frac{1}{2 + \gamma(W_2 - W_1)}\{\gamma + [\delta(f_1 - W_1) + \delta(f_1 - W_2)]\}, \\ \eta_0(f_2) = \frac{1}{2 + \gamma(W_2 - W_1)}\{\gamma + [\delta(f_2 - W_1) + \delta(f_2 - W_2)]\}. \tag{32}$$

The game value is equal to

$$\nu = \frac{2\varphi}{2 + \gamma(W_2 - W_1)}. \tag{33}$$

Let us briefly discuss the physical meaning of the obtained game solution. Since the solutions of Eqs. (18) and (19) usually contain delta-functions, it is more convenient to represent the optimal strategies in the form of a distribution function instead of in the form of probability density. For instance, the distribution function corresponding to the probability density (28) is given by

$$F(f_1) = c_1 1\,[f_1 - W_1] + c_2 (f_1 - W_1) + c_1 1\,[f_1 - W_2], \tag{34}$$

where 1[x] is the unit function, and c_1 and c_2 are constants which, in particular for the payoff function (31), are equal to

$$2c_1 = \frac{2}{2+\gamma(W_2 - W_1)}, \qquad c_2 (W_2 - W_1) = \frac{\gamma(W_2 - W_1)}{2+\gamma(W_2 - W_1)}. \tag{35}$$

Since the γ parameter in (35) determines the "width" of the payoff function (31), we conclude from (35) that, if the payoff function width decreases, the probability portion for the jumps (in other words, the probability that the boundary frequencies W_1 and W_2 will be used) decreases and that the "value" of the carrier frequency uniform variation increases.

Equation (33), which determines the game value, indicates that the noise effectiveness decreases as the $\gamma(W_2 - W_1)$ product value increases. Therefore, in the case of broadband noise, when its spectral density coincides with the payoff function with accuracy to the factor [see (16)], white noise emission is the most effective one for noise.

2. Carrier frequency changes within $[W_1, W_2]$ usually occur in jumps. The consideration of such a game approximation leads to the so-called finite or matrix games, for which the solution methods have been explored to the greatest extent [2].

Let us divide the entire $[W_1, W_2]$ range into $(W_2 - W_1)/h$ subranges. Then

$$f_1 = W_1 + ih, \quad f_2 = W_1 + jh \qquad \left(0 \leqslant i,\ j \leqslant \frac{W_2 - W_1}{h}\right).$$

In this case, the payoff function can be represented by a matrix whose elements a_{ij} are given by

$$a_{ij} = \Phi(|i-j|h) \qquad \left(0 \leqslant i,\ j \leqslant \frac{W_2 - W_1}{h}\right).$$

If the payoff function is different from zero only for

$$|f_1 - f_2| < \varepsilon, \qquad f_1, f_2 \in [W_1, W_2],$$

the game matrix is given by

$$\Phi(|i-j|h) = a_{ij} \neq 0 \quad \text{for} \quad |i-j| < \frac{\varepsilon}{h}, \tag{36}$$

$$\Phi(|i-j|h) \equiv 0 \quad \text{for} \quad |i-j| \geqslant \frac{\varepsilon}{h}. \tag{37}$$

3. By using the S. N. Bernshtein polynomials for approximating* the payoff function, the solution of the game under consideration can be reduced to the solution of the so-called "separable" games.

The payoff function for the separable game can be represented in the following form [2]:

$$\Phi(f_1, f_2) = \sum_{ij} a_{ij} r_i(f_1) s_j(f_2), \tag{38}$$

---

* The following theorem has been proved in the theory of games [2]. If the series $\Phi_1(f_1, f_2), \Phi_2(f_1, f_2), \ldots, \Phi_n(f_1, f_2)$ of payoff functions that are continuous over the square, and which have game values equal to $\nu_1, \nu_2, \ldots, \nu_n$, respectively, is uniformly convergent to the continuous, over-the-square payoff function $\Phi(f_1, f_2)$ with the game value ν, the $\nu_1, \nu_2, \ldots, \nu_n$ series converges to ν.

where a_{ij} are constants, and $r_i(f_1)$ and $s_j(f_2)$ are continuous functions. The general methods for solving separable games are known [2].

S. N. Bernshtein proved [9] that any continuous function $f(x)$ that is defined in the $0 \le x \le 1$ interval can be approximated by polynomials of the following form:

$$f_n(x) = \sum_{m=0}^{n} C_n^m f\left(\frac{m}{n}\right) x^m (1-x)^{n-m}, \tag{39}$$

while $\lim_{n\to\infty} f_n(x) = f(x)$, and C_n^m is the number of combinations of n with respect to m.

As an example, we shall approximate the payoff functions (16) and (17) by S. N. Bernshtein polynomials, assuming that the noise spectral density or, correspondingly, the receiver frequency response, is a rational-fraction function. In this case, the approximating polynomial is given by

$$\Phi_n[(f_1-f_2)^2] = \Phi_n\left[\Delta W^2\left(\frac{f_1-W_1}{\Delta W} - \frac{f_2-W_1}{\Delta W}\right)^2\right] = \Phi_n[\Delta W^2(\nu_1-\nu_2)^2] =$$

$$= \Phi_n(\Delta W^2 \Delta\nu^2) = \sum_{m=0}^{n} \Phi\left(\frac{m}{n}\right) C_n^m \Delta\nu^{2m}(1-\Delta\nu^2)^{n-m}, \tag{40}$$

where $0 \le \nu_1 \le 1$, $0 \le \nu_2 \le 1$.

By applying the binomial theorem to (40), we finally obtain the separable payoff function:

$$\Phi_n(\Delta\nu^2) = \sum_{i=0}^{n} \Delta\nu^{2i} a_i(n) = \sum_{i=0}^{n}\sum_{k=0}^{2i} \nu_1^{2i-k}\nu_2^k a_{ik}(n), \tag{41}$$

where

$$a_i(n) = \sum_{m=0}^{i} (-1)^{i-m}\Phi\left(\frac{m}{n}\right) C_n^m C_{n-m}^{i-m}.$$

The general payoff function (6) can be reduced to the separable payoff function in a similar manner.

Let us consider an example. Assume that

$$\Phi[(\Delta f)^2] = \frac{a^2}{a^2+(\Delta f)^2} = \frac{1}{1+\lambda^2\Delta\nu^2}, \tag{42}$$

where $\lambda^2 = \frac{(W_2-W_1)^2}{a^2}$ is, according to the assumption, a small-value parameter.

The S. N. Bernshtein polynomial for the payoff function (42) will be written thus:

$$\Phi_n[(\Delta\nu)^2] = \sum_{m=0}^{n} \frac{1}{1+\lambda^2\frac{m}{n}} C_n^m \Delta\nu^{2m}(1-\Delta\nu^2)^{n-m} =$$

$$= \sum_{k=0}^{\infty} \frac{\lambda^{2k}}{n^k}(-1)^k \sum_{m=0}^{n} m^k C_n^m \Delta\nu^{2m}(1-\Delta\nu^2)^{n-m}. \tag{43}$$

It can be shown [9] that

$$\sum_{m=0}^{n} m^k C_n^m \Delta\nu^{2m}(1-\Delta\nu^2)^{n-m} = \Delta\nu^{2k} n^k + \Delta\nu^{2(k-1)} n^{k-1} + \dots .$$

Thus, in the limiting case, we have

$$\lim_{n\to\infty} \Phi_n(\Delta\nu^2) = \sum_{k=0}^{\infty}(-1)^k \lambda^{2k}\Delta\nu^{2k}. \tag{44}$$

If the polynomial (43) is a second-degree polynomial (n = 2), for small λ values we obtain

$$\Phi_2[(\Delta\nu)^2] = \Phi(0) + \Delta\nu^2\left[2\Phi\left(\frac{1}{2}\right) - 2\Phi(0)\right] + \Delta\nu^4\left\{\Phi(0) + \Phi(1) - 2\Phi\left(\frac{1}{2}\right)\right\} \approx 1 - \lambda^2\Delta\nu^2. \quad (45)$$

The solution of the game with the payoff function (45) is given by

$$\xi_0(\nu_1) = \delta\left(\nu_1 - \frac{1}{2}\right) = \delta\left(f_1 - \frac{W_1 + W_2}{2}\right),$$

$$\eta_0(\nu_2) = \frac{1}{2}\delta(\nu_2 - 0) + \frac{1}{2}\delta(\nu_2 - 1) = \frac{1}{2}\delta(f_2 - W_1) + \frac{1}{2}\delta(f_2 - W_2). \quad (46)$$

The game value is equal to

$$v = 1 - \frac{\lambda^2}{4}. \quad (47)$$

The exact solution of the game with the payoff function (42) [2] yields the same optimal mixed strategies as the solution obtained from the approximate payoff function (45); however, the game value is equal to

$$v = \frac{1}{1 + \frac{\lambda^2}{4}}. \quad (48)$$

Further, the exact solution (46) is valid for

$$0 \leqslant \lambda^2 \leqslant \frac{4}{3}.$$

Thus, the "approximate" game value is given by the first two terms of the series obtained by expanding the expression for the "exact" game value.

LITERATURE CITED

1. D. Blackwell and M. A. Girshick, Theory of Games and Statistical Decisions [Russian translation] (IL, 1958).
2. I. C. C. McKinsey, Introduction to the Theory of Games (McGraw Hill, New York, 1952).
3. F. M. Woodward, Probability Theory and Information Theory in Radar Applications [in Russian] (Sov. Radio, 1955).
4. V. S. Pugachev, "A method for determining optimal systems with respect to the general Bayes criterion," Izvest. Akad. Nauk SSSR, Otd. Tekh. Nauk, Seriya Énergetiki i Avtomatiki No. 2 (1960).
5. L. A. Vainshtein and V. D. Zubakov, Signal Separation from Random Noise Background [in Russian] (Sov. Radio, 1960).
6. V. S. Pugachev, Theory of Random Functions and Its Application to Automatic Control Problems [in Russian] (Gostekhizdat, 1960) 2nd ed.
7. S. Karlin, "Bell-shaped games," Ann. Math. Studies 3, No. 37 (1957).
8. S. Karlin, "Reduction of the certain classes of games to the integral equations," Ann. Math. Studies 3, No. 28 (1953).
9. V. L. Goncharov, Theory of the Interpolation and Approximation of Functions [in Russian] (Gostekhizdat, 1954).

All abbreviations of periodicals in the above bibliography are letter-by-letter transliterations of the abbreviations as given in the original Russian journal. *Some or all of this periodical literature may well be available in English translation.* A complete list of the cover-to-cover English translations appears at the back of this issue.

CERTAIN NUMERICAL METHODS FOR DETERMINING PERIODICAL MOTIONS OF AUTOMATIC CONTROL SYSTEMS

Yu. I. Neimark

Translated from Avtomatika i Telemekhanika, Vol. 22, No. 1,
pp. 47-56, January, 1961
Original article submitted June 18, 1960

Certain numerical methods for determining periodical motions of automatic control systems are proposed here. For the application of these methods, it is sufficient to know the frequency characteristics of the linear elements and the diagrams of characteristics for the nonlinear elements.

One of the main problems arising in theoretical investigations of the dynamics of automatic control systems is the determination and investigation of periodical motions. As a result of the efforts of many investigators, considerable progress has been made toward solving this problem for systems that are close to linear systems (the small-value parameter method [1-3], the van der Pol method [4-7]), systems that satisfy the autoresonance or filter hypothesis (different variants of the describing function approach [8-11]), sectionally linear systems, the study of which can be reduced to the conversion of a straight line into another straight line [12-15], and for the simplest periodical motions of relay systems [16-19]. However, in many cases encountered in practice, none of these methods can be applied. Moreover, even in those cases where the widely used describing function approach yields good results, there is some doubt about their correctness.

The aim of this paper is to present a numerical method for determining periodical motions of automatic control systems which is based on the method proposed in [20]. The use of this method does not presuppose the knowledge of the analytical equations of motion for the system under investigation. It is sufficient to know the frequency responses of the linear elements, and to have the diagrams of the characteristics of the nonlinear elements.

1. Basic Initial Relationships

Consider an open-loop system consisting of a linear element with the frequency response $K(i\omega) = K_1(\omega) + iK_2(\omega)$, and a nonlinear functional element (Fig. 1). The frequency response $K(i\omega)$ and the nonlinear element characteristic $\Omega(x)$ can be given analytically as well as graphically.

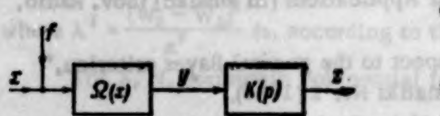


Fig. 1.

The periodical motions will be approximately assigned either in the form of finite segments of a Fourier series, or in the form of a series of values at equidistant instants of time.

If we take into account $\underline{n}$ first harmonics, then, according to [20], between the Fourier coefficients $A_1, A_2, \ldots, A_n, B_0, \ldots, B_n$ of the periodical input quantity

$$x = \sum_{k=0}^{n} A_k \sin k\omega t + B_k \cos k\omega t \tag{1}$$

and the Fourier coefficients of the output quantity periodical component

$$z = \sum_{k=0}^{n} \overline{A}_k \sin k\omega t + \overline{B}_k \cos k\omega t \tag{2}$$

the following relationships hold:

$$\overline{A}_k = \frac{2}{m} \sum_{j=0}^{m-1} \Omega \left[\sum_{s=0}^{n} A_s \sin \frac{2\pi}{m} sj + B_s \cos \frac{2\pi}{m} sj + f_j \right] \times$$

$$\times\left[K_1(k\omega)\sin\frac{2\pi}{m}kj-K_2(k\omega)\cos\frac{2\pi}{m}kj\right],$$

$$B_k=\frac{2}{m}\sum_{j=0}^{m-1}\Omega\left[\sum_{s=0}^{n}A_s\sin\frac{2\pi}{m}sj+B_s\cos\frac{2\pi}{m}sj+f_j\right]\times$$

$$\times\left[K_1(k\omega)\cos\frac{2\pi}{m}kj+K_2(k\omega)\sin\frac{2\pi}{m}kj\right], \tag{3}$$

where $f_0, f_1, \ldots, f_{m-1}$ are the values of the external input quantity $f(t)$, which is applied together with $\underline{x}$ to the nonlinear element input at the instants of time

$$t_0,\ t_1=t_0+\frac{2\pi}{m\omega},\ \ldots,\ t_{m-1}=t_0+\frac{2\pi(m-1)}{m\omega}.$$

Similarly, if $x_0, x_1, \ldots, x_{m-1}$ are the values of the periodical input quantity $\underline{x}$, and $z_0, z_1, \ldots, z_{m-1}$ are the values of the output quantity $\underline{z}$ periodical component at the equidistant instants of time

$$t_0,\ t_1=t_0+\frac{2\pi}{m\omega},\ \ldots,\ t_{m-1}=t_0+\frac{(m-1)2\pi}{m\omega},$$

then

$$z_j=\frac{2}{m}\left\{\sum_{\sigma=0}^{m-1}\Omega(x_\sigma)\sum_{k=0}^{n}\left[K_1(k\omega)\cos\frac{2\pi}{m}k(\sigma-j)+K_2(k\omega)\sin\frac{2\pi}{m}k(\sigma-j)\right]\right\}+f_j. \tag{4}$$

In this, the number $\underline{n}$ of harmonics to be taken into account and the number $\underline{m}$ of equidistant instants of time in the $2\pi/\omega$ period are chosen arbitrarily in correspondence with the required accuracy, but under the condition that $m \geq 2n+1$ must hold. The $\underline{m}$ values equal to 12, 24, and 48 should be preferred if the calculations are to be simplified.

An arbitrary automatic control system with a single nonlinear element can be obtained from the open-loop system (Fig. 1) by feeding the output signal $\underline{z}$ to the input, i.e., by assuming that $z = x$.

Thus, in correspondence with relations (3) and (4), the equations for A_k and B_k or x_j, which approximately determine the periodical motion with the period $2\pi/\omega$ in which we are interested, will be written thus:

$$A_k=\frac{2}{m}\sum_{j=0}^{m-1}\Omega\left[\sum_{s=0}^{n}A_s\sin\frac{2\pi}{m}sj+B_s\cos\frac{2\pi}{m}sj+f_j\right]\times$$

$$\times\left[K_1(k\omega)\sin\frac{2\pi}{m}sj-K_2(k\omega)\cos\frac{2\pi}{m}sj\right],$$

$$B_k=\frac{2}{m}\sum_{j=0}^{m-1}\Omega\left[\sum_{s=0}^{n}A_s\sin\frac{2\pi}{m}sj+B_s\cos\frac{2\pi}{m}sj+f_j\right]\times \tag{5}$$

$$\times\left[K_1(k\omega)\cos\frac{2\pi}{m}sj+K_2(k\omega)\sin\frac{2\pi}{m}sj\right]$$

and, correspondingly, also in the following form:

$$x_j=\frac{2}{m}\left\{\sum_{\sigma=0}^{m-1}\Omega(x_\sigma)\sum_{k=0}^{n}\left[K_1(k\omega)\cos\frac{2\pi}{m}k(\sigma-j)+K_2(k\omega)\sin\frac{2\pi}{m}k(\sigma-j)\right]\right\}+f_j=$$

$$=\frac{2}{m}\sum_{\sigma=0}^{m-1}\Omega(x_\sigma)\,\mathrm{Re}\sum_{s=0}^{n}K(is\omega)\,e^{is(j-\sigma)\frac{2\pi}{m}}+f_j, \tag{6}$$

where, if external disturbances are absent, we put $f_j = 0$.

Equations (5) or (6) make it possible to find in the accepted approximation the periodical motions of the nonlinear system under consideration.

Actually, in the case of forced oscillations, the $2\pi/\omega$ period is known, and Eqs. (5) or (6) constitute a system of 2n + 1 or, correspondingly, m equations with respect to 2n + 1 unknowns $A_1, \ldots, A_n, B_0, \ldots, B_n$ or, respectively, m unknowns $x_0, \ldots, x_{m-1}$. In the case of an autonomous system, i.e., if no external disturbances are present ($f_j = 0$), the period $2\pi/\omega$ is unknown. The missing equation can be obtained by using the arbitrariness of the initial moment t_0 of time measurement, and by choosing it in such a manner that, for instance, $A_1 = 0$ or $B_1 = 0$. For symmetric periodical motions, $x_0 = 0$ can be taken.

The system of equations (6) can be written in the following form:

$$x_j = \sum_{\sigma=0}^{m-1} a_{j\sigma}\,\Omega(x_\sigma) + f_j, \tag{7}$$

where the matrix of coefficients $a_{j\sigma}$ is given by

$$\left\| \begin{matrix} \alpha_0, & \alpha_1, & \alpha_2, \ldots, \alpha_{m-2}, & \alpha_{m-1} \\ \alpha_m, & \alpha_0, & \alpha_1, \ldots, \alpha_{m-3}, & \alpha_{m-2} \\ \alpha_{m-1}, & \alpha_m, & \alpha_0, \ldots, \alpha_{m-4}, & \alpha_{m-3} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_1 & \alpha_2, & \alpha_3, \ldots, \alpha_{m-1}, & \alpha_0 \end{matrix} \right\|, \tag{8}$$

and is, therefore, determined by its first row

$$\alpha_0,\ \alpha_1,\ \alpha_2, \ldots, \alpha_{m-1},$$

where

$$\alpha_\sigma = \frac{2}{m}\sum_{k=0}^{n} K_1(k\omega)\cos\frac{2\pi}{m}k\sigma + K_2(k\omega)\sin\frac{2\pi}{m}k\sigma = \frac{2}{m}\,\mathrm{Re}\sum_{k=0}^{n} K(ik\omega)\,e^{-ik\sigma\frac{2\pi}{m}}. \tag{9}$$

If the nonlinear element characteristic, the external disturbance, and the periodical motion to be found are symmetric, then, for the even value m = 2m', Eq. (6) will be written in the form given by (7) with the following matrix:

$$\left\| \begin{matrix} \alpha_0, & \alpha_1, & \alpha_2, \ldots, \alpha_{m'-2}, & \alpha_{m'-1} \\ -\alpha_{m'-1}, & \alpha_0, & \alpha_1, \ldots, \alpha_{m'-3}, & \alpha_{m'-2} \\ -\alpha_{m'-2}, & -\alpha_{m'-1}, & \alpha_0, \ldots, \alpha_{m'-4}, & \alpha_{m'-3} \\ \cdots & \cdots & \cdots & \cdots \\ -\alpha_1, & -\alpha_2, & -\alpha_3, \ldots, -\alpha_{m'-1}, & \alpha_0 \end{matrix} \right\|, \tag{10}$$

where the values of α_σ are expressed according to Eq. (9) with the substitution of m' for m.

Equations (5) can be written in the following form:

$$A_k + iB_k = \frac{2}{m}K(ik\omega)\sum_{j=0}^{m-1}\Omega(x_j)\,e^{i\frac{2\pi}{m}kj}, \tag{11}$$

where

$$x_j = \sum_{s=0}^{n} A_s\sin\frac{2\pi}{m}sj + B_s\cos\frac{2\pi}{m}sj + f_j = \mathrm{Im}\,\Sigma(A_s + iB_s)\,e^{i\frac{2\pi}{m}sj} + f_j. \tag{12}$$

For symmetric periodical motion and the even m = 2m' value, the even harmonics vanish from Eq. (11), which is then written as

$$A_k + iB_k = \frac{2}{m'}K(ik\omega)\sum_{j=0}^{m'-1}\Omega(x_j)\,e^{i\frac{\pi}{m'}kj}, \tag{11'}$$

where

$$x_j = \mathrm{Im}\sum_{s=1,3,5,\ldots}(A_s + iB_s)\,e^{i\frac{\pi}{m'}sj} + f_j. \tag{12'}$$

2. Method of Trial Hypotheses

If the nonlinear element characteristic is composed of straight-line segments, the solution of Eqs. (5) or (6) can be found in the following manner. A certain assumption (trial hypothesis) is made about the position of the points $x_0, x_1, x_2, \ldots, x_{m-1}$ on the linear sections of the nonlinear element characteristic, after which the system of Eqs. (5) or (6) becomes linear with respect to the A_k and B_k or x_j variables, respectively. This system is used for determining the unknown quantities, and the validity of the initial assumption is verified. If the initial assumption is false, the next hypothesis is set forth, etc.

Example 1.* Determine the forced oscillations of a proportional integral control system under the action of an external disturbance given by $f(t) = 0.5 \sin 2t$, which is fed to the input of a servomotor with the nonlinear characteristic shown in Fig. 2, for the following parameter values:

$$T_a = 5,\ T_k = 0.01,\ \beta = 1,\ T_2^2 = 0.01,\ \delta = 0.1,\ \theta = 1,\ T_i = 100$$

and a number of values of the T_s parameter (T_s^{-1} is the angular coefficient of the middle portion of the nonlinear element characteristic). The control system structural diagram is shown in Fig. 3. In the case under consideration,

$$K(p) = -\frac{(T_i p + 1) + \beta T_i p (T_a p + \theta)(T_2^2 p^2 + T_k p + \delta)}{p(T_a p + \theta)(T_i p + 1)(T_2^2 p^2 + T_k p + \delta)}. \tag{13}$$

The transfer locus $K(i\omega)$ is shown in Fig. 4. We shall use the system (5), assuming that m = 12 and n = 5. By taking into account the symmetry of the periodical motion to be found, this system will be written as

$$x_0 = \frac{1}{3}[0.74\Omega(x_0) + \Omega(x_1) + 0.6\Omega(x_2) + 0.26\Omega(x_3) - 0.2\Omega(x_4) - 0.29\Omega(x_5)],$$

$$x_1 = \frac{1}{3}[0.29\Omega(x_0) + 0.74\Omega(x_1) + \Omega(x_2) + 0.6\Omega(x_3) + 0.26\Omega(x_4) - 0.2\Omega(x_5)] + 0.25,$$

$$x_2 = \frac{1}{3}[0.2\Omega(x_0) + 0.29\Omega(x_1) + 0.74\Omega(x_2) + \Omega(x_3) + 0.6\Omega(x_4) + 0.26\Omega(x_5)] + 0.42, \quad (14)$$

$$x_3 = \frac{1}{3}[-0.26\Omega(x_0) + 0.2\Omega(x_1) + 0.29\Omega(x_2) + 0.74\Omega(x_3) + \Omega(x_4) + 0.6\Omega(x_5)] + 0.5,$$

$$x_4 = \frac{1}{3}[-0.6\Omega(x_0) - 0.26\Omega(x_1) + 0.2\Omega(x_2) + 0.29\Omega(x_3) + 0.74\Omega(x_4) + \Omega(x_5)] + 0.42,$$

$$x_5 = \frac{1}{3}[-\Omega(x_0) - 0.6\Omega(x_1) - 0.26\Omega(x_2) + 0.2\Omega(x_3) + 0.29\Omega(x_4) + 0.74\Omega(x_5)] + 0.25. \tag{14}$$

The remaining values of x_j will be found from the equation $x_j = -x_{j-6}$ (j = 6, 7, 8, 9, 10, 11, 12).

Let us consider the following hypotheses:

a. for $T_s = 0.45$,

$$\Omega(x_0) = \Omega(x_1) = \Omega(x_2) = \Omega(x_3) = 1, \quad \Omega(x_4) = \frac{1}{T_s} x_4, \quad \Omega(x_5) = \frac{1}{T_s} x_5,$$

b. for $T_s = 0.33$,

$$\Omega(x_0) = \Omega(x_1) = \Omega(x_2) = \Omega(x_3) = 1, \quad \Omega(x_4) = \frac{1}{T_s} x_4, \quad \Omega(x_5) = -1,$$

c. for $T_s = 0.1$,

$$\Omega(x_0) = \Omega(x_1) = \Omega(x_2) = \Omega(x_3) = \Omega(x_4) = 1, \quad \Omega(x_5) = -1.$$

After substituting the $\Omega(x_j)$ values in the system of equations (14), we find that, for the hypothesis a,

$$x_0 = 0.95; \quad x_1 = 1.11; \quad x_2 = 1.03; \quad x_3 = 0.58;$$
$$x_4 = -0.15; \quad x_5 = -0.2;$$

* This and the following examples were composed by A. Al'tman and M. Yudkovich.

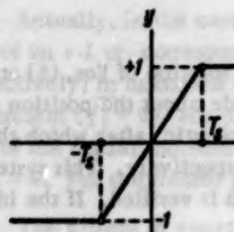


Fig. 2.

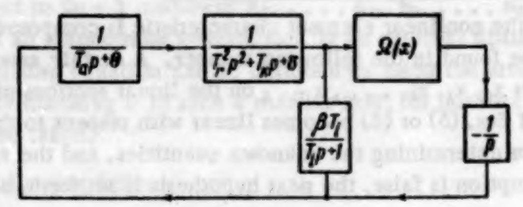


Fig. 3.

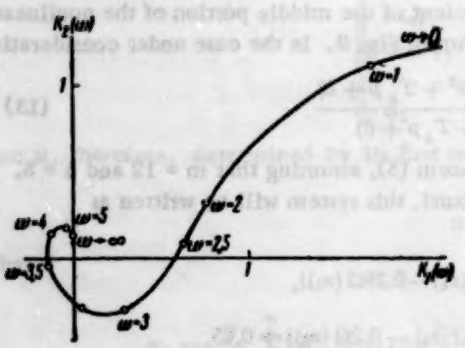


Fig. 4.

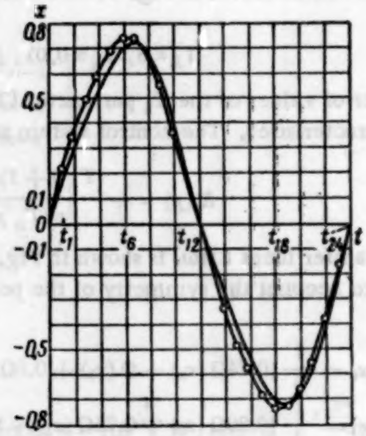


Fig. 5.

for the hypothesis b,

$$x_0 = 0.99;\quad x_1 = 1.16;\quad x_2 = 1.0;\quad x_3 = 0.5;$$
$$x_4 = -0.1;\quad x_5 = -0.5;$$

and for the hypothesis c,

$$x_0 = 0.89;\quad x_1 = 1{,}28;\quad x_2 = 1.27;\quad x_3 = 0.96;$$
$$x_4 = 0.21;\quad x_5 = -0.45.$$

In each of the a, b, and c cases, the hypothesis used proved to be correct (the rejected variants are not given here).

Similar calculations for m = 24 and n = 5 yielded the following results:

a. for $T_s = 0.45$,

$$x_0 = 0.92;\quad x_1 = 1.0;\quad x_2 = 1.12;\quad x_3 = 1.2;\quad x_4 = 0.93;\quad x_5 = 0.76,$$
$$x_6 = 0.5;\quad x_7 = 0.2;\quad x_8 = -0.06;\quad x_9 = -0.24;\ x_{10} = -0.4;\ x_{11} = -0.66;$$

b. for $T_s = 0.33$,

$$x_0 = 1.0;\quad x_1 = 1.07;\quad x_2 = 1.10;\quad x_3 = 1.05;\quad x_4 = 0.92;\quad x_5 = 0.70;$$
$$x_6 = 0{,}4;\quad x_7 = 0.06;\quad x_8 = -0.2;\quad x_9 = -0.39;\ x_{10} = -0.56;\ x_{11} = -0.72;$$

c. for $T_s = 0.1$,

$$x_0 = 0{,}94;\quad x_1 = 1{,}13;\quad x_2 = 1{,}25;\quad x_3 = 1{,}27;\quad x_4 = 1{,}2;\quad x_5 = 1{,}06;$$
$$x_6 = 0{,}86;\quad x_7 = 0{,}5;\quad x_8 = 0{,}11;\quad x_9 = -0{,}2;\quad x_{10} = -0{,}46;\ x_{11} = -0{,}67.$$

E x a m p l e 2. Consider the same proportional integral control system in the absence of external disturbances. The equilibrium state of the system under consideration is stable for $T_s > 0.59$, and it is unstable for $T_s < 0.59$. Let m = 12 and n = 5; then, if we take into account the symmetry of the periodical motion to be determined, the matrix of the system of equations (5) will have the form given by (10), where

$$\begin{aligned}
\alpha_0 &= \frac{1}{3}\,[K_1(\omega) + K_1(3\omega) + K_1(5\omega)],\\
\alpha_1 &= \frac{1}{3}\,[0.85K_1(\omega) - 0.85K_1(5\omega) + 0.5K_2(\omega) + K_2(3\omega) + 0.5K_2(5\omega)],\\
\alpha_2 &= \frac{1}{3}\,[0.5K_1(\omega) - K_1(3\omega) + 0.5K_1(5\omega) + 0.85\,K_2(\omega) - 0.85K_2(5\omega)],\\
\alpha_3 &= \frac{1}{3}\,[K_2(\omega) - K_2(3\omega) + K_2(5\omega)],\\
\alpha_4 &= \frac{1}{3}\,[-0.5K_1(\omega) + K_1(3\omega) - 0.5K_1(5\omega) + 0.85K_2(\omega) - 0.85K_2(5\omega)],\\
\alpha_5 &= \frac{1}{3}\,[-0.85K_1(\omega) + 0.85K_1(5\omega) + 0.5K_2(\omega) + K_2(3\omega) + 0.5K_2(5\omega)].
\end{aligned} \tag{15}$$

In addition to these equations, we shall assume that $x_0 = 0$. Let $T_s = 0.45$ and $T_s = 0.33$. For the first case, we shall assume that

$$\Omega(x_1) = \frac{1}{T_s}\,x_1, \quad \Omega(x_2) = \Omega(x_3) = \Omega(x_4) = 1, \quad \Omega(x_5) = \frac{1}{T_s}\,x_5.$$

In correspondence with this hypothesis, the system of equations for x_1, x_2, x_3, x_4, x_5, and ω will assume the following form:

$$\begin{gathered}
0 = 2.2\alpha_1 x_1 + \alpha_2 + \alpha_3 + \alpha_4 + 2.2\alpha_5 x_5, \quad x_1 = 2.2\alpha_0 x_1 + \alpha_1 + \alpha_2 + \alpha_3 + 2.2\alpha_4 x_5,\\
x_2 = -2.2\alpha_5 x_1 + \alpha_0 + \alpha_1 + \alpha_2 + 2.2\alpha_3 x_5, \quad x_3 = -2.2\alpha_4 x_1 - \alpha_6 + \alpha_0 + \alpha_1 + 2.2\alpha_1 x_5,\\
x_4 = -2.2\alpha_3 x_1 - \alpha_4 - \alpha_5 + \alpha_0 + 2.2\alpha_1 x_5, \quad x_5 = -2.2\alpha_2 x_1 - \alpha_3 - \alpha_4 - \alpha_5 + 2.2\alpha_0 x_5,
\end{gathered}$$

where the α_0, α_1, α_2, α_3, α_4, and α_5 coefficients are given by Eq. (15).

If we express the unknowns x_1 and x_5 from the first and the sixth equation in terms of ω, and substitute them in the second equation, by using the graphical method, we find that $\omega = 2.6$. For this value of ω, we find x_2, x_3, and x_4 from the remaining equations. Thus,

$$x_0 = 0; \quad x_1 = 0.34; \quad x_2 = 0.57; \quad x_3 = 0.68; \quad x_4 = 0.61; \quad x_5 = 0.34$$

and the hypothesis used has been proved correct.

For $T_s = 0.33$, and assuming that

$$\Omega(x_1) = \Omega(x_2) = \Omega(x_3) = \Omega(x_4) = \Omega(x_5) = 1,$$

we obtain the following system:

$$\begin{gathered}
0 = \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5, \quad x_1 = \alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4,\\
x_2 = -\alpha_1 + \alpha_0 + \alpha_1 + \alpha_2 + \alpha_3, \quad x_3 = -\alpha_2 - \alpha_1 + \alpha_0 + \alpha_1 + \alpha_2,\\
x_4 = -\alpha_3 - \alpha_2 - \alpha_1 + \alpha_0 + \alpha_1, \quad x_5 = -\alpha_4 - \alpha_3 - \alpha_2 - \alpha_1 + \alpha_0.
\end{gathered}$$

From the first equation, we find that $\omega = 2.6$, and, from the remaining equations, we find that

$$x_1 = 0.38; \quad x_2 = 0.55; \quad x_3 = 0.73; \quad x_4 = 0.63; \quad x_5 = 0.35.$$

Similar calculations for m = 24 and n = 5 for the case where $T_s = 0.45$ and $T_s = 0.33$ yield, respectively,

$$\begin{aligned}
&\omega = 2.6; && x_1 = 0.15; && x_2 = 0.36; && x_3 = 0.48; && x_4 = 0.57; && x_5 = 0.62;\\
&x_6 = 0.7; && x_7 = 0.69; && x_8 = 0.61; && x_9 = 0.5; && x_{10} = 0.31; && x_{11} = 0.2
\end{aligned}$$

and

$$\begin{aligned}
&\omega = 2.6; && x_1 = 0.2; && x_2 = 0.38;\\
&x_3 = 0.47; && x_4 = 0.59; && x_5 = 0.68;\\
&x_6 = 0.73; && x_7 = 0.72; && x_8 = 0.66;\\
&x_9 = 0.54; && x_{10} = 0.34; && x_{11} = 0.17.
\end{aligned}$$

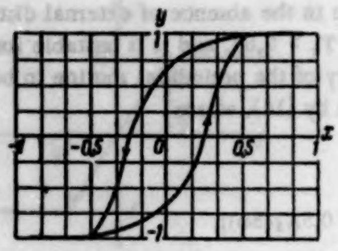


Fig. 6.

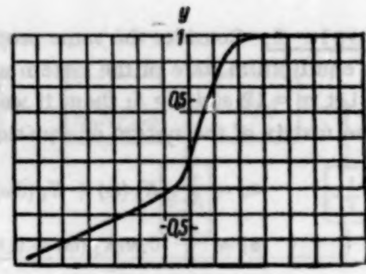


Fig. 7.

For $T_s = 0$, the nonlinear element becomes a relay element, and the periodical regime can be found exactly [16-19]. For the sake of comparison, Fig. 5 shows the diagrams of the exact and the approximate solutions. The approximate solution (curve with white points) was found for m = 24 and n = 5.

3. Method of Consecutive Approximations (Iteration)

The form of equations (5) and (6) is suitable for the direct application of the iteration method. In the case of a nonautonomous system, we take any values of A_j and B_j or x_j for the initial zero approximation and, by substituting them in the right-hand sides of equations (5) or (6), respectively, we find the first approximation; in the same way we find the second approximation by using the first approximation, etc. If the thus obtained consecutive values converge, they converge toward the solution. For an autonomous system, the application of the iteration method is somewhat more complicated; it can be done in the following manner. We write system (6), together with the additional requirement given by

$$x_j = \sum a_{js}(\omega)\, x_s + b_j(\omega), \qquad F(x_0, x_1, \ldots, x_{m-1}; \omega) = 0. \tag{16}$$

For the assigned zero approximation for $x_0, x_1, \ldots, x_{m-1}$ and ω, we find the values of x_j from the first group of equations by using the iteration method in the manner described above; then, by substituting the found values of x_j in the last equation, we find the new value of ω. From the first group of equations we find the next approximation for the x_j variables, etc.

For the A_k and B_k variables, the iteration method can be applied in a similar manner. In the absence of convergence in the iteration of all A_k and B_k variables, it is possible to secure convergence in a number of cases by iteration only with respect to the Fourier coefficients of the higher-order harmonics.

Example 3. We shall determine the self-oscillations for the same control system as in the previous example with the difference, however, that the nonlinear element has the characteristic shown in Fig. 6. For m = 12, n = 5, and $x_0 = 0$, we arrive at the following system:

$$\begin{aligned}
0 &= \alpha_0\Omega(0) + \alpha_1\Omega(x_1) + \alpha_2\Omega(x_2) + \alpha_3\Omega(x_3) + \alpha_4\Omega(x_4) + \alpha_5\Omega(x_5),\\
x_1 &= -\alpha_5\Omega(0) + \alpha_0\Omega(x_1) + \alpha_1\Omega(x_2) + \alpha_2\Omega(x_3) + \alpha_3\Omega(x_4) + \alpha_4\Omega(x_5),\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
x_5 &= -\alpha_1\Omega(0) - \alpha_2\Omega(x_1) - \alpha_3\Omega(x_2) - \alpha_4\Omega(x_3) - \alpha_5\Omega(x_4) + \alpha_0\Omega(x_5),
\end{aligned}$$

where values of α_0 to α_5 are determined by Eq. (15). For the initial approximation, we shall take $x_1 = x_2 = \ldots = x_5 = 0.5$. Then, from the first equation, we find $\omega = 2.6$ and, from the remaining equations,

$$x_1 = 0.33; \quad x_2 = 0.57; \quad x_3 = 0.72; \quad x_4 = 0.67; \quad x_5 = 0.38.$$

For the new values of $x_1, x_2, \ldots, x_5$, we find $\omega = 2.62$ from the first equation and, from the remaining five equations, we find $x_1 = 0.31$, $x_2 = 0.51$, $x_3 = 0.67$, $x_4 = 0.61$, and $x_5 = 0.35$.

The third approximation yields $x_1 = 0.31$, $x_2 = 0.54$, $x_3 = 0.66$, $x_4 = 0.60$, and $x_5 = 0.35$.

The obtained solution can be rendered more accurate by using m = 24 and n = 5, and by taking the just-found values for the initial approximation. The intermediate values are found by linear interpolation. After two iterations, we obtain: $x_1 = 0.13$, $x_2 = 0.28$, $x_3 = 0.41$, $x_4 = 0.5$, $x_5 = 0.58$, $x_6 = 0.61$, $x_7 = 0.61$, $x_8 = 0.54$, $x_9 = 0.46$, $x_{10} = 0.33$, and $x_{11} = 0.20$.

Example 4. Assume that the nonlinear element has the asymmetric characteristic shown in Fig. 7. In system (5), we shall put m = 24 and n = 5, and we shall additionally assume that $B_1 = 0$. Since $K(0) = \infty$, $B_0 = 0$. By taking $A_1 = 0.34$, $A_2 = A_3 = \ldots = A_5 = 0$, $B_0 = \ldots = B_5 = 0$ for the initial (zero) approximation, and by using Eqs. (12') and the graph in Fig. 7, we find, in the order of increasing values of the number j, the values of $\Omega(x_j)$: 0, 0.3, 0.53, 0.72, 0.83, 0.88, 0.9, 0.88, 0.83, 0.72, 0.53, 0.3, 0.0, −0.18, −0.24, −0.28, −0.31, −0.33, −0.34, −0.33, −0.31, −0.28, −0.24, and −0.18.

By substituting these values of $\Omega(x_j)$ in the equation $B_1 = 0$, we find that $\omega = 2.6$. After substituting the found value of ω in Eq. (5), we find the following approximation for the A_k and B_k Fourier coefficients: $A_1 = 0.39$, $A_2 = 0.02$, $A_3 = A_4 = A_5 = 0$; $B_0 = B_1 = 0$, $B_2 = 0.01$, $B_3 = B_4 = B_5 = 0$.

The third approximation is a repetition of the foregoing with the accepted accuracy; in this, $\omega = 2.62$, $A_1 = 0.42$, $A_2 = 0.02$, $A_3 = A_4 = A_5 = 0$; $B_0 = B_1 = 0$, $B_2 = 0.01$, $B_3 = 0.01$, $B_4 = B_5 = 0$.

4. Method of Extension with Respect to a Parameter

The extension of the solution with respect to a parameter is a very flexible and efficient method for determining the solutions of equations (5) or (6). In this, the parameter with respect to which the solution is extended can be either one of the parameters of the system under consideration or a parameter artificially introduced for extending the solutions from equations with known solutions to equations whose solutions are to be determined.

The combination of the above-described methods with the extension with respect to a parameter is convenient, since the required hypothesis can always be readily found, or a good initial approximation is already available.

A more independent application of the method of extension with respect to a parameter can also be used. In the general case, if we assume that $K(i\omega)$ and $\Omega(x)$ are functions of a parameter μ, and if we differentiate the system of equations (5) or (6) with respect to this parameter, we obtain the following system of differential equations:

$$\frac{dx_j}{d\mu} = \frac{m}{2}\left\{\sum_{\sigma=0}^{m-1}\left(\Omega'_\mu(x_\sigma) + \Omega'_x\frac{dx_\sigma}{d\mu}\right)\sum_{s=0}^{n}\operatorname{Re}K(is\omega)\,e^{is\frac{2\pi}{m}(j-\sigma)} + \right.$$
$$\left. + \sum_{\sigma=0}^{m-1}\Omega(x_\sigma)\sum_{s=0}^{n}\operatorname{Re}\left[isK'_p(is\omega)\frac{d\omega}{d\mu} + K'_\mu(is\omega)\right]e^{is\frac{2\pi}{m}(j-\sigma)}\right\} \quad (17)$$

or, correspondingly,

$$\frac{dA_k}{d\mu} + i\frac{dB_k}{d\mu} = \frac{2}{m}\left\{K(ki\omega)\sum_{j=0}^{m-1}\left(\Omega'_\mu(x_j) + \Omega'_x(x_j)\frac{dx_j}{d\mu}\right)e^{i\frac{2\pi}{m}kj} + \right.$$
$$\left. + \left[ikK'_p(ik\omega)\frac{d\omega}{d\mu} + K'_\mu(ik\omega)\right]\sum_{j=0}^{m-1}\Omega(x_j)\,e^{i\frac{2\pi}{m}kj}\right\}, \quad (18)$$

where

$$\frac{dx_j}{d\mu} = \sum_{s=0}^{n}\frac{dA_s}{d\mu}\sin\frac{2\pi}{m}sj + \frac{dB_s}{d\mu}\cos\frac{2\pi}{m}sj.$$

Assume that, for $\mu = \mu_0$, the solution of the system of equations (5) or (6) is known. By using this solution for the initial conditions for $\mu = \mu_0$, and by solving the system of differential equations (17) or (18) by one of the approximate methods, we find the necessary solution for a certain interval of the parameter μ values.

It should be noted that, in the case where the nonlinear element has a sectionally linear characteristic for all variables except ω, the systems of equations (17) and (18) represent systems of linear differential equations with constant coefficients for each extension interval where no changes occur in the arrangement of $x_0, x_1, \ldots, x_{m-1}$ values on the linear portions of the nonlinear characteristic.

In conclusion, it should be emphasized that the application of the method of extension with respect to a parameter is connected with studies of possible bifurcations of periodic motions [21].

LITERATURE CITED

1. A. A. Andronov and A. A. Vitt, "On the mathematical theory of self-oscillating systems with two degrees of freedom," Collected works of A. A. Andronov [in Russian] (AN SSSR Press, 1956).
2. B. V. Bulgakov, Oscillations [in Russian] (Gostekhizdat, 1954).
3. I. G. Malkin, Application of Lyapunov and Poincaré Methods in the Theory of Nonlinear Oscillations [in Russian] (Gostekhizdat, 1949).
4. A. A. Andronov and A. A. Vitt, "On the van der Pol's engagement theory," Collected Works of A. A. Andronov [in Russian] (AN SSSR Press, 1956).
5. N. V. Butenin, "On one Calvin problem in the theory of clocks," Zhur. Éksp. i Teoret. Fiz. 10, 11 (1940).
6. N. V. Butenin, "Self-oscillations of autopilot stands," Tr. LKVVIA No. 3 (1943).
7. N. N. Bogolyubov and Yu. A. Mitropol'skii, Application of Asymptotic Methods in the Theory of Nonlinear Oscillations [in Russian] (Fizmatgiz, 1958).
8. N. M. Krylov and N. N. Bogolyubov, Introduction to Nonlinear Mechanics [in Russian] (AN UkrSSR Press, 1937).
9. L. S. Gol'dfarb, "On certain nonlinearities in control systems," Avtomatika i Telemekhanika 8, No. 5 (1947).
10. G. K. Krug, "Certain problems in the nonlinearity effect on servomotor characteristics," Transaction of the All-Union Conference on Automatic Control Theory [in Russian] (AN SSSR Press, 1955) Vol. 1.
11. M. A. Aizerman and I. M. Smirnova, "On the application of small-value parameter methods in the investigation of periodic regimes in automatic control systems without small-value parameters," Collection Devoted to A. A. Andronov [in Russian] (AN SSSR Press, 1955).
12. A. A. Andronov, A. A. Vitt, and S. É. Khaikin, Theory of Oscillations [in Russian] (Fizmatgiz, 1959).
13. N. A. Fufaev, "Electromagnetic interrupter theory," Collection Devoted to A. A. Andronov [in Russian] (AN SSSR Press, 1955).
14. A. S. Alekseev, "A two-position temperature regulator with advance range," Collection Devoted to A. A. Andronov [in Russian] (AN SSSR Press, 1955).
15. Yu. I. Neimark, Yu. K. Maklakov, and L. P. Elkina, "Circulation of pulses in a heavily nonlinear system with dispersive lagging feedback," Radiotekhnika i Élektronika 3, 11 (1958).
16. A. I. Lur'e, Some Nonlinear Problems in Automatic Control Theory [in Russian] (Gostekhizdat, 1951).
17. Ya. Z. Tsypkin, Theory of Automatic Control Relay Systems [in Russian] (Gostekhizdat, 1955).
18. Yu. I. Neimark, "On the stability and periodic regimes in relay systems," Avtomatika i Telemekhanika 14, No. 5 (1953).
19. Yu. I. Neimark, "On sliding conditions and periodic motions in relay systems," Uch. Zap. Radiofiz. Fakul'teta Gor'kovskogo Gos. Univ. 30 (1956).
20. Yu. I. Neimark, "A numerical method for determining periodic motions," Izvest. Vyssh. Uchebn. Zaved., Radiofizika 2, No. 6 (1959).
21. Yu. I. Neimark, "Application of the point reflection method in nonlinear oscillation theory," Izvest. Vyssh. Uchebn. Zaved., Radiofizika 1, No. 5, 6 (1958).

All abbreviations of periodicals in the above bibliography are letter-by-letter transliterations of the abbreviations as given in the original Russian journal. ***Some or all of this periodical literature may well be available in English translation.*** **A complete list of the cover-to-cover English translations appears at the back of this issue.**

DETERMINATION OF AIRCRAFT CONTROL EQUATIONS FOR THE OPTIMUM PATH UNDER VARIABLE-WIND FLIGHT CONDITIONS

V. D. Matytsin

Translated from Avtomatika i Telemekhanika, Vol. 22, No. 1,
pp. 57-66, January, 1961
Original article submitted April 20, 1960

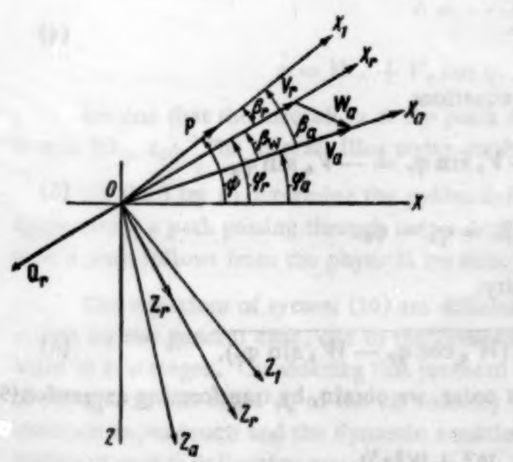
The present article is concerned with the extremal problem in determining the conditions necessary for the quickest aircraft flight from one point in space to another, under conditions where the wind changes with respect to coordinates and in time. The minimum flight time is taken as the criterion for the optimum path. Variation calculus is used as the mathematical tool (the relative extremum variation problem).

If the effect of wind along the flight path is given detailed consideration, the advantage of the shortest-distance flight becomes doubtful. It is sometimes more advantageous somewhat to lengthen the route if, by taking advantage of the wind, the over-all flight duration can be reduced.

The determination of the optimum flight paths is of practical interest from the point of minimizing the flight time. The knowledge of such flight paths makes it possible to obtain more complete information on the maximum aircraft range, and to perform the planned flight with minimum fuel consumption [1].

At the present time, cruising flights are conducted at altitudes where the wind velocity can attain considerable magnitudes. Consequently, the equations of motion must reflect, to the fullest possible extent, the effect of wind on the aircraft dynamics.

In connection with the fact that flights take place practically at the same altitude, we shall consider the aircraft motion problem in the course plane. The coordinate systems and the basic notation (see figure), which are used in correspondence with the accepted standard [2], are given below.



The following coordinate systems are used: XOZ is the terrestrial coordinate system, X_1OZ_1 is the bound coordinate system, and X_rOZ_r is the velocity coordinate system.

In the X_aOZ_a coordinate system, the OX_a axis has the direction of the flight velocity vector.

In composing the equations of motion, the following notation was used: V_a is the path (absolute) aircraft velocity, V_r is the aircraft air (relative) velocity, W is the wind velocity, P is the aircraft propelling force, Q_r is the front drag force, Z_r is the lateral force, ψ is the aircraft yaw angle, φ_r is the direction of the air velocity vector, φ_a is the aircraft course angle (direction of the path velocity vector), β_r is the relative (aerodynamic) slip angle, which is formed by the positive directions of the OX_r and OX_1 axes, β_a is the absolute slip angle, which is formed by the positive directions of the OX_a and OX_1 axes, and β_W is the additional slip angle due to the wind, which is formed by the positive directions of the OX_a and OX_r axes.

The aerodynamic forces and moments depend on changes in the air (relative) velocity; they are given by the following equations:

$$Q_r = C_x \frac{\rho S}{2} V_r^2,$$

$$Z_r = C_z \frac{\rho S}{2} V_r^2 = Z_r^{\beta_r} \beta_r, \tag{1}$$

$$M_y = m_y \frac{\rho S}{2} l V_r^2 = \frac{\rho S l}{2} V_r^2 \left[m_y^{\beta_r} \beta_r + m_y^{\delta} \delta + \frac{l}{2V_r} m_y^{\omega_y} \dot{\psi} \right].$$

Under actual flight conditions, the wind exerts the most frequent and the longest effects on the aircraft.

The wind can be defined in a terrestrial coordinate system by its projections*:

$$W = \{W_x(x, z, t), \quad W_z(x, z, t)\}.$$

In the aircraft motion equations, we shall take into account the fact that the wind changes the aerodynamic forces acting on the aircraft by changing the aircraft velocity with respect to air, without, however, affecting the inertial terms of the equation of motion.

The inertial terms depend on the terrestrial (absolute) velocity and acceleration, and the aerodynamic forces depend on the air (relative) velocity.

S. B. Puzrin was the first to derive the aircraft motion equations by taking into account the wind effect [3]. We shall use his method.

Let us project all forces acting on the aircraft on the OX_a and OZ_a axes. Projection on the OX_a axis yields

$$m \frac{dV_a}{dt} = P \cos \beta_a - Q_r \cos \beta_w + Z_r \sin \beta_w, \tag{2}$$

and projection on the OZ_a axis gives us

$$\frac{mV_a^2}{R_a} = mV_a \frac{d\varphi_a}{dt} = -P \sin \beta_a + Q_r \sin \beta_w + Z_r \cos \beta_w, \tag{3}$$

where R_a is the instantaneous curvature radius of the absolute path.

The equation of moments with respect to the y axis has the following form:

$$I_y \frac{d^2\psi}{dt^2} = M_y. \tag{4}$$

The system of Eqs. (1)-(3) is supplemented by the kinematic equations

$$\dot{x} = W_x + V_r \cos \varphi_r = V_a \cos \varphi_a, \quad \dot{z} = W_z - V_r \sin \varphi_r = -V_a \sin \varphi_a,$$
$$\psi = \beta_r + \varphi_r = \beta_a + \varphi_a, \qquad \beta_w = \beta_a - \beta_r = \varphi_r - \varphi_a. \tag{5}$$

By using system (5), we find the expression for the path velocity:

$$V_a = \sqrt{\dot{x}^2 + \dot{z}^2} = \sqrt{W_x^2 + W_z^2 + V_r^2 + 2V_r (W_x \cos \varphi_r - W_z \sin \varphi_r)}. \tag{6}$$

Considering W_x/V_r and W_z/V_r as small quantities of the first order, we obtain, by transforming expression (6),

$$V_a = V_r \left(1 + 2\frac{W_x \cos \varphi_r - W_z \sin \varphi_r}{V_r} + \frac{W_x^2 + W_z^2}{V_r^2}\right)^{1/2} \approx$$

$$\approx V_r \left(1 + \frac{W_x \cos \varphi_r - W_z \sin \varphi_r}{V_r}\right).$$

***Flight in the horizontal plane is considered.**

Thus,

$$V_a = V_r + W_x \cos\varphi_r - W_z \sin\varphi_r. \tag{7}$$

System (5) also makes it possible to determine the additional slip angle β_W, which is due to the wind:

$$\begin{aligned} V_a \cos(\varphi_r - \beta_W) &= W_x + V_r \cos\varphi_r = V_a \cos\varphi_r \cos\beta_W + V_a \sin\varphi_r \sin\beta_W \\ -V_a \sin(\varphi_r - \beta_W) &= W_z - V_r \sin\varphi_r = -V_a \sin\varphi_r \cos\beta_W + V_a \cos\varphi_r \sin\beta_W \end{aligned} \tag{8}$$

By multiplying the first equation in system (8) by $\sin\varphi_r$, and the second equation by $\cos\varphi_r$, we obtain, after addition:

$$V_a \sin\beta_W = W_x \sin\varphi_r + W_z \cos\varphi_r$$

$$\sin\beta_W = \frac{W_x \sin\varphi_r + W_z \cos\varphi_r}{V_a}. \tag{9}$$

By taking into account (1)-(7), and the smallness of angles β_W and β_r, we have

$$V_a = V_r + W_x \cos\varphi_r - W_z \sin\varphi_r,$$

$$mV_a\left(\frac{d\varphi_r}{dt} - \frac{d\beta_W}{dt}\right) = -P(\beta_r + \beta_W) + Q_r\beta_W + Z_r^{\beta_r}\beta_r.$$

$$I_y \frac{d^2\psi}{dt^2} = Sl\frac{\rho V_r^2}{2}\left[m_y^{\beta_r}\beta_r + m_y^{\delta}\delta\right] + Sl^2\frac{\rho}{4}V_r m_y^{\omega_y}\dot{\psi}.$$

By replacing the equation of motion about the center of gravity by the balancing relation, we finally obtain

$$\begin{gathered} V_a = V_r + W_x \cos\varphi_r - W_z \sin\varphi_r, \\ mV_a\left(\frac{d\varphi_r}{dt} - \frac{d\beta_W}{dt}\right) = (-P + Z_r^{\beta_r})\beta_r + (-P + Q_r)\beta_W, \\ \delta = -\frac{m_y^{\beta_r}}{m_y^{\delta}}\beta_r, \qquad \beta_W = \frac{W_x \sin\varphi_r + W_z \cos\varphi_r}{V_a}, \\ \dot{x} = W_x + V_r \cos\varphi_r, \quad \dot{z} = W_z - V_r \sin\varphi_r, \quad \psi = \varphi_r + \beta_r. \end{gathered} \tag{10}$$

Assume that the aircraft is at the point $A(x_0, z_0)$ at the initial moment of time, and that the flight destination is $B(x_1, z_1)$. The aircraft flies under conditions where the wind velocity depends on coordinates and the time.

We shall try to determine the rudder deflection equation, which will make it possible to complete the flight along a path passing through points $A(x_0, z_0)$ and $B(x_1, z_1)$ in the shortest possible time. The existence of such a path follows from the physical meaning of the problem.

The equations of system (10) are differential linkage equations. The analysis obviously cannot be completed for the general case, due to the complexity of the linkage equations. Therefore, we shall solve this problem in two stages. Considering this problem as a purely kinematic problem, we shall determine the equation of changes in the angle φ_r of the air velocity vector in order to obtain the optimum path. Then, by using the obtained dependence and the dynamic equations, and by taking into account the wind, we shall determine the optimum rudder deflection equation.

Considering that the air velocity vector value does not change considerably in flight, we shall consider V_r as a constant quantity. For these conditions, the problem was solved by I. M. Gyunter [4], who reduced it to the so-called Meyer variation problem. We shall now give a simpler and shorter derivation of the equation of changes in the air velocity vector angle for the optimum path.

We shall formally consider the flight time as a quantity depending on a certain parameter ξ, $t = t(\xi)$; then $x = x(\xi)$, $z = z(\xi)$; at the initial instant of time,

$$t_0 = t(\xi_0), \quad x_0 = x(\xi_0), \quad z_0 = z(\xi_0),$$

and at the final instant of time,

$$t_1 = t(\xi_1), \quad x_1 = x(\xi_1), \quad z_1 = z(\xi_1). \tag{11}$$

The time differential will be given by $dt = \frac{dt(\xi)}{d\xi} d\xi$, and the flight time will be determined by the integral

$$t = \int_{\xi_0}^{\xi_1} \frac{dt(\xi)}{d\xi} d\xi. \tag{12}$$

The kinematic linkage equations will assume the following form:

$$\frac{dx}{d\xi} = \frac{dt}{d\xi}(W_x + V_r \cos\varphi_r), \quad \frac{dz}{d\xi} = \frac{dt}{d\xi}(W_z - V_r \sin\varphi_r). \tag{13}$$

Following M. A. Lavrent'ev and L. A. Lyusternik [5], we shall denote the $dt/d\xi$ function by F, i.e.,

$$\frac{dt}{d\xi} = F[\xi, x(\xi), z(\xi), \varphi_r(\xi), \dot{x}_\xi, \dot{z}_\xi, \dot{\varphi}_{r\xi}].$$

Equation (12) will assume the following form:

$$t = \int_{\xi_0}^{\xi_1} F[\xi, x, z, \varphi_r, \dot{x}_\xi, \dot{z}_\xi, \dot{\varphi}_{r\xi}] d\xi, \tag{14}$$

and the linkage equations (13) will become

$$\frac{dx}{d\xi} = (W_x + V_r \cos\varphi_r) F, \quad \frac{dz}{d\xi} = (W_z - V_r \sin\varphi_r) F. \tag{15}$$

The flight time (14) can be considered as a functional of F. The F function itself will be determined from the first equation of system (15), namely

$$F = \frac{\dot{x}_\xi}{W_x + V_r \cos\varphi_r}. \tag{16}$$

The flight path in which we are interested must satisfy the linkage equation which will be obtained by eliminating the F function from (15):

$$\Phi(\xi, x, z, \varphi_r, \dot{x}_\xi, \dot{z}_\xi) = \dot{z}_\xi - \dot{x}_\xi \frac{W_z - V_r \sin\varphi_r}{W_x + V_r \cos\varphi_r} = 0. \tag{17}$$

The class of the acceptable paths to be determined will consist of curves satisfying the differential linkage equation (17) and the boundary conditions (11). We shall now find the condition for the extremum of functional (14). For the solution of this problem, we shall use the following theorem.

The x, z, and φ_r functions that secure the extremum of the functional

$$t = \int_{\xi_0}^{\xi_1} F(\xi, x, z, \varphi_r, \dot{x}_\xi, \dot{z}_\xi, \dot{\varphi}_{r\xi}) d\xi$$

in the presence of the linkage equation

$$\Phi(\xi, x, z, \varphi_r, \dot{x}_\xi, \dot{z}_\xi, \dot{\varphi}_{r\xi}) = 0$$

satisfy for a suitable choice of the $\lambda(\xi)$ factor the Euler equations composed for the functional

$$t^* = \int_{\xi_0}^{\xi_1} [F + \lambda(\xi)\Phi] d\xi = \int_{\xi_0}^{\xi_1} H d\xi.$$

The $\lambda(\xi)$, x, z, and φ_r functions are determined from the following Euler equations:

$$\frac{\partial H}{\partial x}-\frac{d}{d\xi}\frac{\partial H}{\partial \dot{x}_\xi}=0,\qquad \frac{\partial H}{\partial z}-\frac{d}{d\xi}\frac{\partial H}{\partial \dot{z}_\xi}=0,\qquad \frac{\partial H}{\partial \varphi_r}-\frac{d}{d\xi}\frac{\partial H}{\partial \dot{\varphi}_{r\xi}}=0 \tag{18}$$

and from the linkage equation*

$$\Phi=\dot{z}_\xi-\dot{x}_\xi\frac{W_z-V_r\sin\varphi_r}{W_x+V_r\cos\varphi_r}.$$

In our case, the system of equations (18) assumes the following form:

$$\begin{gathered}\frac{\partial F}{\partial x}+\lambda\frac{\partial\Phi}{\partial x}-\frac{d}{d\xi}\left[\frac{\partial F}{\partial \dot{x}_\xi}+\lambda\frac{\partial\Phi}{\partial \dot{x}_\xi}\right]=0,\\ \frac{\partial F}{\partial z}+\lambda\frac{\partial\Phi}{\partial z}-\frac{d}{d\xi}\left[\frac{\partial F}{\partial \dot{z}_\xi}+\lambda\frac{\partial\Phi}{\partial \dot{z}_\xi}\right]=0,\\ \frac{\partial F}{\partial \varphi_r}+\lambda\frac{\partial\Phi}{\partial \varphi_r}-\frac{d}{d\xi}\left[\frac{\partial F}{\partial \dot{\varphi}_{r\xi}}+\lambda\frac{\partial\Phi}{\partial \dot{\varphi}_{r\xi}}\right]=0,\\ \Phi=\dot{z}_\xi-\dot{x}_\xi\frac{W_z-V_r\sin\varphi_r}{W_x+V_r\cos\varphi_r}=0.\end{gathered}\tag{19}$$

It is essential that λ be a function of ξ only.

By using expression (16) for F, we shall determine the $\partial F/\partial x$, $\partial F/\partial z$, $\partial F/\partial\varphi_r$, $\partial F/\partial\dot{x}_\xi$, $\partial F/\partial\dot{z}_\xi$, and $\partial F/\partial\dot{\varphi}_{r\xi}$ derivatives:

$$\begin{gathered}\frac{\partial F}{\partial x}=-\dot{x}_\xi\frac{\frac{\partial W_x}{\partial x}-V_r\sin\varphi_r\frac{\partial\varphi_r}{\partial x}}{(W_x+V_r\cos\varphi_r)^2}=-F\frac{\frac{\partial W_x}{\partial x}-V_r\sin\varphi_r\frac{\partial\varphi_r}{\partial x}}{W_x+V_r\cos\varphi_r},\\ \frac{\partial F}{\partial z}=-\dot{x}_\xi\frac{\frac{\partial W_x}{\partial z}-V_r\sin\varphi_r\frac{\partial\varphi_r}{\partial z}}{(W_x+V_r\cos\varphi_r)^2}=-F\frac{\frac{\partial W_x}{\partial z}-V_r\sin\varphi_r\frac{\partial\varphi_r}{\partial z}}{W_x+V_r\cos\varphi_r},\\ \frac{\partial F}{\partial\varphi_r}=-\dot{x}_\xi\frac{-V_r\sin\varphi_r}{(W_x+V_r\cos\varphi_r)^2}=F\frac{V_r\sin\varphi_r}{W_x+V_r\cos\varphi_r},\\ \frac{\partial F}{\partial\dot{x}_\xi}=\frac{1}{W_x+V_r\cos\varphi_r},\qquad \frac{\partial F}{\partial\dot{z}_\xi}=0,\qquad \frac{\partial F}{\partial\dot{\varphi}_{r\xi}}=0.\end{gathered}\tag{20}$$

Similarly, by using the linkage equation (17), we obtain

$$\begin{gathered}\frac{\partial\Phi}{\partial x}=F\frac{W_z-V_r\sin\varphi_r}{W_x+V_r\cos\varphi_r}\left(\frac{\partial W_x}{\partial x}-V_r\sin\varphi_r\frac{\partial\varphi_r}{\partial x}\right)-F\left(\frac{\partial W_z}{\partial x}-V_r\cos\varphi_r\frac{\partial\varphi_r}{\partial x}\right),\\ \frac{\partial\Phi}{\partial z}=F\frac{W_z-V_r\sin\varphi_r}{W_x+V_r\cos\varphi_r}\left(\frac{\partial W_x}{\partial z}-V_r\sin\varphi_r\frac{\partial\varphi_r}{\partial z}\right)-F\left(\frac{\partial W_z}{\partial z}-V_r\cos\varphi_r\frac{\partial\varphi_r}{\partial z}\right),\\ \frac{\partial\Phi}{\partial\varphi_r}=-FV_r\sin\varphi_r\frac{W_z-V_r\sin\varphi_r}{W_x+V_r\cos\varphi_r}+FV_r\cos\varphi_r,\\ \frac{\partial\Phi}{\partial\dot{x}_\xi}=-\frac{W_z-V_r\sin\varphi_r}{W_x+V_r\cos\varphi_r},\qquad \frac{\partial\Phi}{\partial\dot{z}_\xi}=1,\qquad \frac{\partial\Phi}{\partial\dot{\varphi}_\xi}=0.\end{gathered}\tag{21}$$

By taking into account (7), the third equation of (21) can be reduced to the form

$$\frac{\partial\Phi}{\partial\varphi_r}=F\frac{V_a V_r}{W_x+V_r\cos\varphi_r}.$$

*See [5], Ch. V, paragraph 22; also [6], Ch. IV, paragraph 2.

Since all the necessary mathematical operations for determining the optimum law of φ_r variation have been performed, we can put $t = \xi$ in systems (20) and (21). Then, $\dot{x}_\xi = \dot{x}_t$, $\dot{z}_\xi = \dot{z}_t$, $\dot{\varphi}_{r\xi} = \dot{\varphi}_{rt}$. In this case, it follows from (10) and (16) that $F = 1$.

For $F = 1$ and by taking into account (15), (20), and (21), system (19) will have the following form:

$$\begin{gathered}
\frac{\frac{\partial W_x}{\partial x} - V_r \sin\varphi_r \frac{\partial \varphi_r}{\partial x}}{W_x + V_r \cos\varphi_r} + \lambda \frac{W_z - V_r \sin\varphi_r}{W_x + V_r \cos\varphi_r}\left(\frac{\partial W_x}{\partial x} - V_r \sin\varphi_r \frac{\partial \varphi_r}{\partial x}\right) - \\
- \lambda\left(\frac{\partial W_z}{\partial x} - V_r \cos\varphi_r \frac{\partial \varphi_r}{\partial x}\right) - \frac{d}{dt}\left[\frac{1}{W_x + V_r \cos\varphi_r} - \lambda \frac{W_z - V_r \sin\varphi_r}{W_x + V_r \cos\varphi_r}\right] = 0, \\
\frac{\frac{\partial W_x}{\partial z} - V_r \sin\varphi_r \frac{\partial \varphi_r}{\partial z}}{W_x + V_r \cos\varphi_r} + \lambda \frac{W_z - V_r \sin\varphi_r}{W_x + V_r \cos\varphi_r}\left(\frac{\partial W_x}{\partial z} - V_r \sin\varphi_r \frac{\partial \varphi_r}{\partial z}\right) - \\
- \lambda\left(\frac{\partial W_z}{\partial z} - V_r \sin\varphi_r \frac{\partial \varphi_r}{\partial z}\right) - \frac{d\lambda}{dt} = 0, \\
\frac{V_r \sin\varphi_r}{W_x + V_r \cos\varphi_r} + \lambda \frac{V_a V_r}{W_x + V_r \cos\varphi_r} = 0, \\
\dot{x} = W_x + V_r \cos\varphi_r, \qquad \dot{z} = W_z - V_r \sin\varphi_r.
\end{gathered} \tag{22}$$

We shall determine λ from the third equation of system (22):

$$\lambda = -\frac{\sin\varphi_r}{V_a}. \tag{23}$$

The first equations of system (22) can be written as

$$\begin{gathered}
-\left(\frac{\partial W_x}{\partial x} - V_r \sin\varphi_r \frac{\partial \varphi_r}{\partial x}\right)\left[\frac{1}{W_x + V_r \cos\varphi_r} - \lambda \frac{W_z - V_r \sin\varphi_r}{W_x + V_r \cos\varphi_r}\right] - \\
- \lambda\left(\frac{\partial W_z}{\partial x} - V_r \cos\varphi_r \frac{\partial \varphi_r}{\partial x}\right) - \frac{d}{dt}\left[\frac{1}{W_x + V_r \cos\varphi_r} - \lambda \frac{W_z - V_r \sin\varphi_r}{W_x + V_r \cos\varphi_r}\right] = 0, \\
-\left(\frac{\partial W_x}{\partial z} - V_r \sin\varphi_r \frac{\partial \varphi_r}{\partial z}\right)\left[\frac{1}{W_x + V_r \cos\varphi_r} - \lambda \frac{W_z - V_r \sin\varphi_r}{W_x + V_r \cos\varphi_r}\right] - \\
- \lambda\left(\frac{\partial W_z}{\partial z} - V_r \cos\varphi_r \frac{\partial \varphi_r}{\partial z}\right) - \frac{d\lambda}{dt} = 0.
\end{gathered} \tag{24}$$

By using (7) and (23), we find that

$$\begin{gathered}
\frac{1}{W_x + V_r \cos\varphi_r} - \lambda \frac{W_z - V_r \sin\varphi_r}{W_x + V_r \cos\varphi_r} = \frac{V_r + W_x \cos\varphi_r - V_r \sin^2\varphi_r}{V_a (W_x + V_r \cos\varphi_r)} = \\
= \frac{V_r \cos^2\varphi_r + W_x \cos\varphi_r}{V_a (W_x + V_r \cos\varphi_r)} = \frac{\cos\varphi_r}{V_a}.
\end{gathered} \tag{25}$$

System (24) assumes the final form:

$$\begin{gathered}
-\frac{d}{dt}\left(\frac{\cos\varphi_r}{V_a}\right) + \\
+ \frac{-\frac{\partial W_x}{\partial x}\cos\varphi_r + V_r \sin\varphi_r \cos\varphi_r \frac{\partial \varphi_r}{\partial x} + \frac{\partial W_z}{\partial x}\sin\varphi_r - V_r \sin\varphi_r \cos\varphi_r \frac{\partial \varphi_r}{\partial x}}{V_a} = 0, \\
-\frac{d}{dt}\left(\frac{-\sin\varphi_r}{V_a}\right) + \\
+ \frac{-\frac{\partial W_x}{\partial z}\cos\varphi_r + V_r \sin\varphi_r \cos\varphi_r \frac{\partial \varphi_r}{\partial z} + \frac{\partial W_z}{\partial z}\sin\varphi_r - V_r \sin\varphi_r \cos\varphi_r \frac{\partial \varphi_r}{\partial z}}{V_a} = 0,
\end{gathered}$$

or

$$\frac{d}{dt}\frac{\cos\varphi_r}{V_a}+\frac{\frac{\partial W_x}{\partial x}\cos\varphi_r-\frac{\partial W_z}{\partial x}\sin\varphi_r}{V_a}=0,$$

$$\frac{d}{dt}\frac{\sin\varphi_r}{V_a}+\frac{\frac{\partial W_z}{\partial z}\sin\varphi_r-\frac{\partial W_x}{\partial z}\cos\varphi_r}{V_a}=0,$$

or

$$\cos\varphi_r\frac{d}{dt}\frac{1}{V_a}+\frac{-\sin\varphi_r\dot{\varphi}_r+\frac{\partial W_x}{\partial x}\cos\varphi_r-\frac{\partial W_z}{\partial x}\sin\varphi_r}{V_a}=0,$$

$$\sin\varphi_r\frac{d}{dt}\frac{1}{V_a}+\frac{\cos\varphi_r\dot{\varphi}_r+\frac{\partial W_z}{\partial z}\sin\varphi_r-\frac{\partial W_x}{\partial z}\cos\varphi_r}{V_a}=0. \tag{26}$$

By multiplying the first of the equations (26) by $\sin\varphi_r$, the second by $\cos\varphi_r$, and by subtracting the first from the second, we obtain

$$\frac{d\varphi_r}{dt}+\left(\frac{\partial W_z}{\partial z}-\frac{\partial W_x}{\partial x}\right)\sin\varphi_r\cos\varphi_r+\frac{\partial W_z}{\partial x}\sin^2\varphi_r-\frac{\partial W_x}{\partial z}\cos^2\varphi_r=0. \tag{27}$$

It should be noted that the difference between the signs in Eq. (27) and those in the results obtained in [4] can be explained by the different choice of the $\underline{z}$ axis positive direction and, correspondingly, the positive direction of the W_z wind component. The coordinate axes used in the present paper are those recommended by the standard [2].

The above derivation is shorter, which is due to the simpler form of the functional under investigation,

$$F=\frac{\dot{x}_\xi}{W_x+V_r\cos\varphi_r},$$

in comparison with

$$[V^2+W_x^2+W_y^2]F^2+2[W_x\dot{x}_\xi+W_y\dot{y}_\xi]F-(\dot{x}_\xi^2+\dot{y}_\xi^2)=0,$$

that was used in [4].

Thus, over the optimum path, the air velocity vector changes its direction according to Eq. (27).

Finally, for determining the rudder deflection equation, we have the system of Eqs. (10) and (27):

$$V_a=V_r+W_x\cos\varphi_r-W_z\sin\varphi_r,\quad \beta_W=\frac{W_z\cos\varphi_r+W_x\sin\varphi_r}{V_a},$$

$$mV_a(\dot{\varphi}_r-\dot{\beta}_W)=(-P+Z_r^{\beta_r})\beta_r+(-P+Q_r)\beta_W,\quad \delta=-\frac{m_y^{\beta_r}}{m_y^{\delta}}\beta_r,$$

$$\dot{\varphi}_r+\left(\frac{\partial W_z}{\partial z}-\frac{\partial W_x}{\partial x}\right)\sin\varphi_r\cos\varphi_r+\frac{\partial W_z}{\partial x}\sin^2\varphi_r-\frac{\partial W_x}{\partial z}\cos^2\varphi_r=0, \tag{28}$$

$$\dot{x}=W_x+V_r\cos\varphi_r,\quad \dot{z}=W_z-V_r\sin\varphi_r.$$

We shall find $\dot{\beta}_W$ from the second of the equations (28):

$$\dot{\beta}_W=\frac{W_xV_a\sin\varphi_r+W_xV_a\cos\varphi_r\dot{\varphi}_r+\dot{W}_zV_a\cos\varphi_r-W_zV_a\sin\varphi_r\dot{\varphi}_r}{V_a^2}-$$

$$-\frac{W_x\dot{V}_a\sin\varphi_r+W_z\dot{V}_a\cos\varphi_r}{V_a^2},$$

or, by expanding $\dot V_a$, we obtain

$$\dot\beta_W = \frac{W_x\cos\varphi_r - W_z\sin\varphi_r}{V_a}\dot\varphi_r + \frac{\dot W_x V_a\sin\varphi_r + \dot W_z V_a\cos\varphi_r}{V_a^2} -$$

$$- \frac{W_x\sin\varphi_r(\dot W_x\cos\varphi_r - W_x\sin\varphi_r\dot\varphi_r - \dot W_z\sin\varphi_r - W_z\cos\varphi_r\dot\varphi_r)}{V_a^2} -$$

$$- \frac{W_z\cos\varphi_r(\dot W_x\cos\varphi_r - W_x\sin\varphi_r\dot\varphi_r - \dot W_z\sin\varphi_r - W_z\cos\varphi_r\dot\varphi_r)}{V_a^2}.$$

By using Eq. (7),

$$W_x\cos\varphi_r - W_z\sin\varphi_r = V_a - V_r;$$

after regrouping, we obtain

$$\dot\beta_W = \frac{(V_a - V_r)V_a + (W_x\sin\varphi_r + W_z\cos\varphi_r)^2}{V_a^2}\dot\varphi_r +$$

$$+ \frac{\dot W_x(V_a\sin\varphi_r - W_x\cos\varphi_r\sin\varphi_r - W_z\cos^2\varphi_r)}{V_a^2} +$$

$$+ \frac{\dot W_z(V_a\cos\varphi_r + W_x\sin^2\varphi_r + W_z\sin\varphi_r\cos\varphi_r)}{V_a^2}$$

or, by substituting $V_a = V_r + W_x\cos\varphi_r - W_z\sin\varphi_r$,

$$\dot\beta_W = \frac{V_a(V_a - V_r) + (W_x\sin\varphi_r + W_z\cos\varphi_r)^2}{V_a^2}\dot\varphi_r +$$

$$+ \frac{\dot W_x(V_r\sin\varphi_r + W_x\cos\varphi_r\sin\varphi_r - W_z\sin^2\varphi_r - W_x\sin\varphi_r\cos\varphi_r - W_z\cos^2\varphi_r)}{V_a^2} +$$

$$+ \frac{\dot W_z(V_r\cos\varphi_r + W_x\cos^2\varphi_r - W_z\sin\varphi_r\cos\varphi_r + W_x\sin^2\varphi_r + W_z\sin\varphi_r\cos\varphi_r)}{V_a^2}.$$

Neglecting the higher-order terms, W_x^2/V_r^2, W_z^2/V_r^2, W_xW_z/V_r^2, after regrouping, we have

$$\dot\beta_W = \frac{V_a - V_r}{V_a}\dot\varphi_r + \frac{-\dot z\dot W_x + \dot x\dot W_z}{V_a^2} \qquad (29)$$

or

$$-\dot\beta_W + \dot\varphi_r \simeq \frac{V_r}{V_a}\dot\varphi_r + \frac{\dot z\dot W_x - \dot x\dot W_z}{V_a^2}. \qquad (30)$$

By using the second and the third equation of system (28), as well as (30), we obtain

$$\frac{V_r}{V_a}\dot\varphi_r + \frac{\dot z\dot W_x - \dot x\dot W_z}{V_a^2} = \frac{-P + Z_r^{\beta_r}}{mV_a}\beta_r + \frac{(-P + Q_r)(W_x\sin\varphi_r + W_z\cos\varphi_r)}{mV_a^2}$$

or

$$\dot\varphi_r = \frac{\dot W_z\dot z - \dot W_x\dot x}{V_a V_r} + \frac{(-P + Z_r^{\beta_r})\beta_r}{mV_r} - \frac{(-P + Q_r)(W_x\sin\varphi_r + W_z\cos\varphi_r)}{mV_a V_r}.$$

By taking into account the fourth and the fifth equation of system (28), the equation for determining the optimum rudder deflection law will have the following form:

$$\left(\frac{\partial W_x}{\partial x} - \frac{\partial W_z}{\partial z}\right)\sin\varphi_r\cos\varphi_r + \frac{\partial W_x}{\partial z}\cos^2\varphi_r - \frac{\partial W_z}{\partial x}\sin^2\varphi_r =$$

$$= \frac{\dot{W}_z(W_x + V_r\cos\varphi_r) - \dot{W}_x(W_z - V_r\sin\varphi_r)}{V_a V_r} - \frac{m_y^\delta}{m_y^{\beta_r}}\frac{-P + Z_r^{\beta_r}}{mV_r}\delta +$$

$$+ \frac{(-P + Q_r)(W_x\sin\varphi_r + W_z\cos\varphi_r)}{mV_r V_a}.$$

By expanding W_x and W_z, after regrouping, we finally obtain the rudder deflection equation:

$$\delta = \frac{m_y^{\beta_r}}{m_y^\delta}\frac{-P + Q_r}{-P + Z_r^{\beta_r}}\frac{W_x\sin\varphi_r + W_z\cos\varphi_r}{V_a} +$$

$$+ \frac{m_y^{\beta_r}}{m_y^\delta}\frac{m}{(-P + Z_r^{\beta_r})V_a}\left\{\frac{\partial W_z}{\partial t}(W_x + V_r\cos\varphi_r) - \frac{\partial W_x}{\partial t}(W_z - V_r\sin\varphi_r) +\right.$$

$$+ \left(\frac{\partial W_z}{\partial z} - \frac{\partial W_x}{\partial x}\right)(W_x W_z + W_z V_r\cos^3\varphi_r - V_r W_x\sin^3\varphi_r) -$$

$$- \frac{\partial W_x}{\partial z}(V_r^2 + W_z^2 + V_r W_x\cos^3\varphi_r + V_r W_z\sin^3\varphi_r - 3V_r W_z\sin\varphi_r) +$$

$$\left. + \frac{\partial W_z}{\partial x}(V_r^2 + W_x^2 - V_r W_z\sin^3\varphi_r - V_r W_x\cos^3\varphi_r + 3V_r W_x\cos\varphi_r)\right\},$$

where φ_r and V_a satisfy the equations

$$\varphi_r = \left(\frac{\partial W_x}{\partial x} - \frac{\partial W_z}{\partial z}\right)\sin\varphi_r\cos\varphi_r + \frac{\partial W_x}{\partial z}\cos^2\varphi_r - \frac{\partial W_z}{\partial x}\sin^2\varphi_r,$$

$$V_a = V_r + W_x\cos\varphi_r - W_z\sin\varphi_r.$$

The optimum path coordinates are determined by the equations

$$\dot{x} = W_x + V_r\cos\varphi_r, \quad \dot{z} = W_z - V_r\sin\varphi_r.$$

LITERATURE CITED

1. G. F. Molokanov, Calculation of the Wind Effect in Long-Range Flight [in Russian] (Voenizdat, 1957).
2. Coordinate Axes and the Basic Notation Used in Aerodynamic and Hydrodynamic Calculations in Aircraft Design [in Russian] (Standartgiz, 1953).
3. S. B. Puzrin, "Differential equations of aircraft motion under disturbance conditions in the longitudinal symmetry plane in the 'terrestrial' coordinate system," Transactions of the Seminar on Motion Stability Theory, ed. by N. D. Moiseev, Tr. VVIA 4 (1948).
4. I. M. Gyunter, Variation Calculus Course [in Russian] (Gostekhizdat, 1941).
5. M. A. Lavrent'ev and L. A. Lyusternik, Variation Calculus Course [in Russian] (Gostekhizdat, 1950).
6. L. É. Él'sgol'ts, Variation Calculus [in Russian] (Gostekhizdat, 1958).

All abbreviations of periodicals in the above bibliography are letter-by-letter transliterations of the abbreviations as given in the original Russian journal. *Some or all of this periodical literature may well be available in English translation.* A complete list of the cover-to-cover English translations appears at the back of this issue.

AUTOMATIC OPTIMIZATION OF SPATIAL DISTRIBUTION.* I

L. N. Fitsner
Translated from Avtómatika i Telemekhanika, Vol. 22, No. 1,
pp. 67-76, January, 1961
Original article submitted July 22, 1960

The principles of automating the spatial distribution of some physical substance are examined. Certain theoretical questions concerned with the solution of the stated problem by means of automatic search are discussed.

1. Class of Solvable Problems

It becomes necessary in solving some technical problems to distribute some physical substance Φ in three-dimensional space in such a manner that the values of Φ will approximate some requirement at $\underline{n}$ given points in the space in the best manner.

For example, let it be necessary to distribute the material allowed for machining parts of complex shape from blanks. To do this, the blank is moved about in space so that the individual points on its surface approximate the corresponding points in a space given beforehand in the best manner. Here the distance between the points on the surface of the blank and the corresponding ones in the space is optimized. The problem of distributing flows in space is quite close to the given problem in the method of its solution.

Let us assume that there are several sound sources located at $\underline{m}$ points in the space, and that it is necessary to find the power of each of the sources in such a way that the sound energy at n other points approximates certain required values in the best manner.

Analogous problems may arise in distributing the flows of other physical things, such as light, radio waves, etc.

The principles of automation discussed below can also be applied in automatic parts assembiage, in the pattern layout of materials for stamping, etc.

We notice certain peculiarities in the class of technical problems being examined. The values of Φ at individual points in the space are interrelated and, as a rule, cannot be set independently. The number of points $\underline{m}$ at which the values of Φ are given, and the number of mechanisms $\underline{n}$ which act on Φ, are not equal in the general case. Ordinarily, it is not possible to get values of Φ which precisely agree with requirements if a limited number of points are available; as a result, the problem solution is only an optimum approximation from one point of view.

Fig. 1.

It is assumed that the required distribution law does not change with time, and that the problem is considered solved when the best approximation of Φ to the given distribution is attained. Deviations of Φ from the required values can be measured at all $\underline{m}$ points of the given Φ.

All the results of the study presented below are related only to one problem, important from a technicoeconomic view, and concerned with the distribution of the material allowed for machining parts of complex shape from blanks. However, a large part of the results obtained are applicable also to other problems of this class.

2. Distribution of Cut Material

In machining the blanks of parts on a machine tool, it is important to position them in the best manner, so that the layer of material removed will be approximately identical at all places on the surface being machined.

*This article was reported at the IAT of the Academy of Sciences, USSR, in the Session on the Theory of Optimum Self-adjusting and Mutually Interconnected Control Systems, Automatic Control Theory Seminar, May 11, 1960.

This will result in the smallest amount of material to be cut away, thus permitting a reduction in the machining time and amount of waste. If the blank has a complex shape, it is extremely difficult, as a rule, to chuck it in the tool; many auxiliary attachments are necessary, and even with their aid painstaking labor is required. Therefore, it is important to automate this process.

The outline of a blank (solid line) and outline of the part (broken line) are shown in Fig. 1 (for simplicity, a flat part will be considered). Let the blank be moved without deformation in three degrees of freedom (progressive movement along x_1 and x_2, and rotation x_3 around an axis perpendicular to the plane of the figure). The problem is to vary x_1, x_2, and x_3 so that the best, from the viewpoint of some criterion, mutual disposition of the broken and solid outlines will be achieved.

Let us assume that removing a uniform layer of cut material results in an ideal part. We will call such a blank outline an isoform outline of the ideal part. In this case, the blank can be set so that the normals AB and A_1B_1, erected at any two points on the ideal part outline, will be equal to one another. Such a mutual disposition will be ideal. However, in general, the blank outline is not an isoform of the ideal part, and the problem can only be one of finding the best approximation to the ideal case, from the standpoint of some criterion or other.

Let us consider three criteria of approximation having practical interest.

Let us designate the distance (AB) between the outlines by ξ_i, and assume that measurements are made at m points along the part outline. We will assume that ξ_i is always positive. Then the first criterion will be

$$Q_1 = \max_{(i)} \xi_i = \max \xi_i (x_1, x_2, x_3) \qquad (i = 1, 2, \ldots, m). \tag{1}$$

The optimum positioning problem consists of choosing x_1, x_2, and x_3 so that Q_1 will be minimized:

$$\min Q_1 = \min_{(x_1, x_2, x_3)} \{\max \xi_i (x_1, x_2, x_3)\} \qquad (i = 1, 2, \ldots, m). \tag{2}$$

In the second criterion, the minima of ξ_i are maximized. This kind of blank positioning eliminates the dangerous condition of having negative cut material:

$$Q_2 = \min_{(i)} \xi_i = \min \xi_i (x_1, x_2, x_3) \qquad (i = 1, \ldots, m). \tag{3}$$

It is necessary to find x_1, x_2, and x_3 which maximize Q_2:

$$\max_{(x_1, x_2, x_3)} Q_2 = \max_{(x_1, x_2, x_3)} \{\min \xi_i (x_1, x_2, x_3)\} \qquad (i = 1, \ldots, m). \tag{4}$$

Finally, the third criterion may be the dispersion of ξ_i, i.e., the difference between the maximum and minimum:

$$Q_3 = Q_1 - Q_2 = \max \xi_i (x_1, x_2, x_3) - \min \xi_j (x_1, x_2, x_3) \qquad (i = 1, \ldots, m), \quad (j = 1, \ldots, m). \tag{5}$$

If this criterion is used, the optimum positioning problem consists of minimizing it:

$$\min_{(x_1, x_2, x_3)} Q_3 = \min_{(x_1, x_2, x_3)} \{\max \xi_i (x_1, x_2, x_3) - \min \xi_j (x_1, x_2, x_3)\} \qquad (i = 1, \ldots, m), \quad (j = 1, \ldots, m). \tag{6}$$

If the part blank is isoform, all three criteria result in a uniform cut. All the considerations presented relate to the case where three measurements are made.

3. Automation Principles

The machining process can be automated with three types of automatic control systems: systems without feedback, automatic control systems, and automatic search systems. We will deal with each of these systems.

Figure 2 shows two isoform parts: one by a solid line (A), and the other by a broken line (B). The distance between the sides of pentagons A and B can be measured by the indicators ξ_1-ξ_5, and part A can be moved along the axes x_1 and x_2 by systems without feedback. The displacement along the axes mentioned can be found from the following formulas:

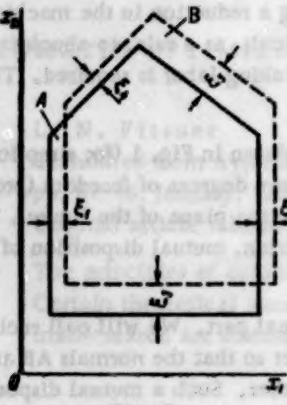


Fig. 2.

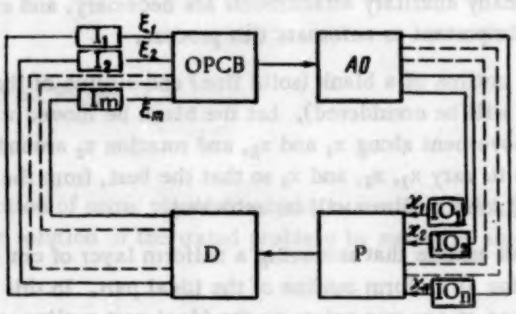


Fig. 3.

$$\Delta x_1 = \xi_1 - \delta, \quad \Delta x_2 = \xi_3 - \delta. \tag{7}$$

It is assumed here that the size of cut material δ is known beforehand. This limitation is not necessary if two automatic control systems are used for positioning, the first of which moves the part along x_1 into a position where

$$\xi_1 - \xi_2 = 0, \tag{8}$$

and the second determines the position along axis x_2:

$$\frac{\xi_4 + \xi_5}{2} - \xi_3 = 0. \tag{9}$$

Let us examine a more general case of optimizing the positioning of isoform parts. If the position deviations from the desired optimum are small, and if the functions

$$\xi_i(x_1, x_2, \ldots, x_n) \quad (i = 1, \ldots, m) \tag{10}$$

have continuous partial first derivatives, the following set of equations is justified:

$$\begin{aligned}
\xi_1 &= \frac{\partial \xi_1}{\partial x_1} x_1 + \frac{\partial \xi_1}{\partial x_2} x_2 + \ldots + \frac{\partial \xi_1}{\partial x_n} x_n, \\
\xi_2 &= \frac{\partial \xi_2}{\partial x_1} x_1 + \frac{\partial \xi_2}{\partial x_2} x_2 + \ldots + \frac{\partial \xi_2}{\partial x_n} x_n, \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
\xi_m &= \frac{\partial \xi_m}{\partial x_1} x_1 + \frac{\partial \xi_m}{\partial x_2} x_2 + \ldots + \frac{\partial \xi_m}{\partial x_n} x_n.
\end{aligned} \tag{11}$$

The coefficients of $\partial \xi_i / \partial x_j$ ($i = 1, \ldots, m$; $j = 1, \ldots, n$) are assumed to be constant and can be determined beforehand experimentally for the given set of equations. The set of algebraic equations (11) can be solved by one of the known methods on a computer [1], and the values of x_j found can be used to govern the actuating mechanisms of the system by means of $\underline{n}$ systems without feedback, or $\underline{n}$ completely independent automatic control systems. Data on the cut material size is not needed if this method is used; isoformity of the parts and stability of the coefficients of $\partial \xi_i / \partial x_j$ in time are necessary.

Automatic search is the best general method for solving the given problem, and is useful for the case of optimizing the machining of nonisoform parts. A block diagram of a system for optimizing this process by the automatic search method is shown in Fig. 3. Platform P with part D fastened on it is moved in $\underline{n}$ degrees of freedom by actuating mechanisms $IO_1, IO_2, \ldots, IO_n$. The deviation of individual points on the blank from a position given by a template is measured by indicators $I_1, I_2, \ldots, I_m$. The outputs of the latter act on the input of the optimum positioning criterion block (OPCB), which generates a signal Q in correspondence with one of the criteria (1), (3), or (5). The automatic optimizer AO excites the actuating mechanisms ($IO_1, IO_2, \ldots, IO_n$), and the minimums of Q_1 or Q_3, or the maximum of Q_2 is found by means of automatic search.

4. Theoretical Questions Related to the Automation of the Spatial Positioning

A number of theoretical questions arise in solving problems related to the optimization of spatial positioning by the automatic search method. The main ones are the following: a) choosing the method of automatic search; b) evaluating the effect of the system inertial elements, and also the static characteristic of the search object on the stability and accuracy of the search process; c) evaluating the effect of backlash and stick-slip friction; d) evaluating the effect of random disturbances; e) evaluating the effect of the insensitive zones of the optimizer; f) finding the optimum form of trial probes; g) synthesizing a coordinate converter which decreases the search time; h) finding the optimum scan amplitude and period; and j) finding a search law which is optimum in use of time.

The greater part of these problems can be studied analytically. However, it is not desirable to use analytic methods for investigating some of these problems, since it may be necessary to develop special apparatus to synthesize automatic search systems automatically; these devices are called automatic synthesizers.

The theoretical studies of optimum positioning of automatic search systems which have been enumerated, done both analytically and with the aid of automatic synthesizers, and also a description of the operating principle of the synthesizer itself, will be the subject of succeeding articles.

5. Concerning the Choice of the Automatic Search Method

It is difficult to base the choise of search method for the optimum positioning of blanks during machining on a strictly mathematical analysis, and we will therefore limit ourselves only to general engineering considerations.

The accuracy and time of search, and the complexity of the actual control apparatus required, are ordinarily the basic choice criteria. A number of the specific peculiarities of the system must be considered in choosing the search method. The platform with the attached part, and the actuating mechanisms may have significant inertial moments which, in general, will not be identical for each of the $\underline{n}$ degrees of freedom.

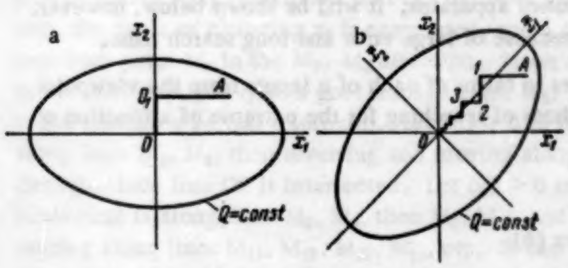


Fig. 4.

Backlash and friction will be present in the motion-transmitting mechanisms; inertial and resistance moments may depend greatly on the deflection angles of the actuating mechanisms. Random interferences may affect the search system; these are mainly caused by surface unevenness of the mounted blank.

The known methods of automatic search in terms of many variables can be divided [2] into blind search methods (scanning, random search), and methods using an analysis of intermediate results (gradient and steepest slope, cyclic search of particular extremes [3]). The scanning and random search methods are associated with a long search time, and thus can be recommended only for the case when there are several extremes of the quantity being optimized, $Q(x_1, \ldots, x_n)$, and organized search methods, which assume that Q has only one extreme, are inapplicable. The gradient and steepest slope methods require that not only the sign, but the magnitude of $\partial Q/\partial x_i$ be determined beforehand. Because of the peculiarities of the systems mentioned, large error is incurred in determining the magnitudes of the partial derivatives, and thus these methods lose their basic advantages. Moreover, the latter require additional time for trial probes, and comparatively complex apparatus. The gradient method, in which each of $\underline{n}$ inputs is excited simultaneously by sinusoidal signals of different frequency [4], requires that the components of the output signal Q, caused by each of the inputs, be determined. In general, a platform with six degrees of freedom means that the response Q to six input signals must be determined. Due to nonlinearities in the system, which cause higher-order harmonics, and because the inertial parameters of the elements are not constant, analysis of the output signal becomes practically impossible.

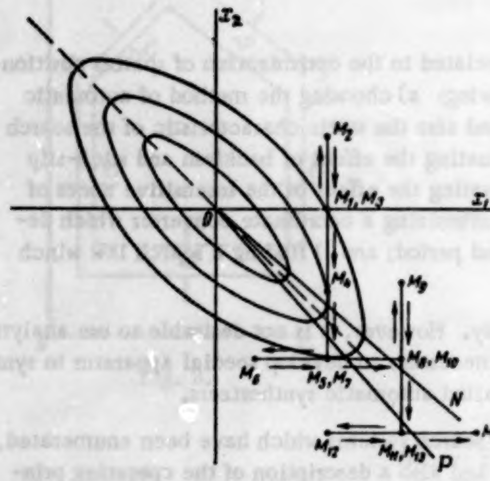


Fig. 5.

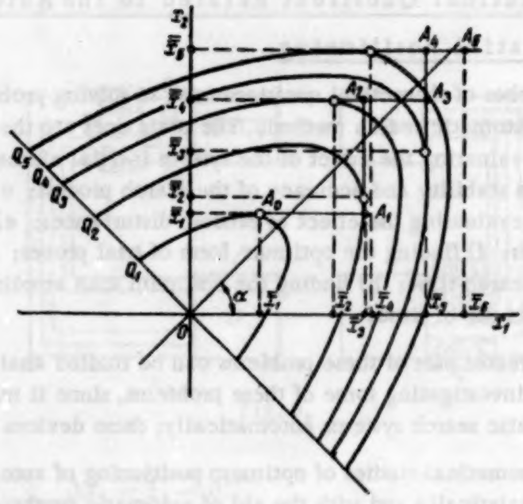


Fig. 6.

The method of cyclic search of partial extremes in terms of each of n inputs is the most suitable method for determining optimum blank positioning (Gauss – Zaidel method). This method, in combination with the method given in [6], allows one to combine trial and working movements. For each of the variables in this case, only the first step can be made in an incorrect direction, and when the direction of movement is changed, backlash must be taken up only once. Since the method requires that only the sign of $\partial Q / \partial x_i$ be found, the peculiarities of the search system do not have as strong an effect as in the methods which have been examined. The given method permits one to accomplish search with very simple control apparatus. It will be shown below, however, that this method, without additional perfection, is not useful because of large error and long search time.

Let us evaluate methods of searching for partial extremes in terms of each of n inputs from the viewpoint of the indicated criteria. We will examine the following methods of searching for the extreme of a function of one variable:

1) self-oscillation method [5];

2) with memory of the peak values of the extreme index [5];

3) with continuous change of the inputs, and measurement of the output increments at equal time intervals [6];

4) step method with constant time intervals between steps [5];

5) step method with variable time intervals between steps [6];

6) step method with constant step amplitude [5];

7) step method with variable step amplitude [7];

8) extrapolation methods [8].

Methods 1, 7, and 8 require that the magnitude of $\partial Q / \partial x_i$ be determined, and, as a result, have a number of deficiencies which were mentioned above. Method 2 is not useful because of the effect of interference; individual peaks and valleys may be taken for peak values when the blank is moved. False change of the search direction results, and this leads to an impermissible length of search time. The advantages of method 5 cannot be used in the given system, since large time intervals between steps near the extreme lengthen the search process. Method 3 is the most suitable for the problem being solved, and also method 4 in combination with method 6. We note that the optimum action form is comparatively complex in methods 4 and 6, due to the presence of inertial elements and restrictions on the coordinates. The problem of determining it will be discussed in the next article. The methods given permit one to solve problems with comparatively simple control apparatus.

6. Effect of the Static Characteristic of the Search Object on Accuracy and Search Time

The search method which involves finding partial extremes cyclically depends essentially on the distribution of the surface on which the extreme points lie relative to the coordinate axes. Thus, when the minimum of the function

$$Q = \frac{x_1^2}{a} + \frac{x_2^2}{b} \tag{12}$$

is found by cyclic search along axes x_1 and x_2, the system moves to the desired point O after only one partial extreme (O_1) is found (Fig. 4a). If the symmetry axis of an elliptical paraboloid (12) is rotated with respect to the search axes, the search is substantially lengthened (Fig. 4b). The system proceeds consecutively from point A through the partial extreme points 1, 2, 3, . . .

Self-oscillations near a point lying to one side of the extreme are possible in such a system at the end of the search process. On the x_1, x_2 plane, let the Q = const lines have the shape of "elongated" ellipses, whose symmetry axes do not coincide with the search axes (Fig. 5). Let us examine the fourth quadrant of the plane. The geometric locus of points at which the tangent to the ellipse Q = const is vertical is the straight line ON; the geometric locus of points at which the tangent to the ellipse Q = const is horizontal is the line OP. Therefore, in moving along a vertical line (for example M_2, M_5), the partial minimum of Q is reached at the intersection of this vertical line with ON, and in moving along a horizontal line (for example M_6, M_{10}), the partial minimum of Q is reached at the intersection of this horizontal line with OP.

Let us assume that the system moves in steps from M_1 as a starting point. At first, x_2 changes, and the system takes a step to point M_2, for example. Since the increment $\Delta Q > 0$ in this case, a reversal occurs, and the next step is taken in the opposite direction; the system proceeds to point M_3, which coincides with M_1. The system simultaneously remembers that ΔQ has been greater than zero once. If the same sign of ΔQ occurs in another step, the cycle of changing x_2 is completed, and a cycle of changing x_1 is begun. The system takes the next step from point M_3 in the M_3, M_4 direction. Since $\Delta Q < 0$ here, subsequent movement is in the same direction. In the next step, the system goes through point M_5. However, line ON is intersected in this case, and the system passes through a partial minimum. Let $\Delta Q > 0$ in this case. Then searching is done along axis x_1, moving at first along lines M_5, M_6, then reversing and moving along lines M_6, M_7, and then M_7, M_8. A partial minimum is passed through, since line OP is intersected. Let $\Delta Q > 0$ in the next step. Then searching is again done along axis x_2; movement is along lines M_8, M_9, then M_9, M_{10}, and M_{10}, M_{11}. The search is once more changed to the x_1 axis, moving along lines M_{11}, M_{12}, M_{13}, M_{14}, etc. It can be seen from Fig. 5 that the mapping point of the system moves away from the true minimum point, and may move away a considerable amount, as will be shown below. One must use a method which includes an alternation in the direction of the first step in the cycle so that the "staircase" M_1, M_5, M_8, M_{11}, . . . will not lead the mapping point far from the minimum O.

In Fig. 5, the first steps in the cycles for the variable x_2 are always taken in a positive direction – these are steps M_1, M_2; M_8, M_9, etc. The first steps in the cycles for the variable x_1 are always made in a negative direction – these are steps M_5, M_6; M_{11}, M_{12}, etc. However, this is not obligatory. The first step in the first cycle of the x_2 variable may be made in a positive direction, the first step in the second cycle in a negative direction, first step in the third cycle positive again, etc. The alternation of signs in the first steps is analogous for x_1. One can then see, by plotting a graph like that of Fig. 5, that the mapping point of the system will move in the vicinity of the minimum O, and will not get far from there. The generalization of this principle to any number of variables is evident. We will show that the deviation of the point around which self-oscillations occur from the extreme point may become significant if first-step sign alternation is not used.

The lines Q = const of an elliptical paraboloid (12) are shown in Fig. 6. Let searching be started along the x_1 variable from point $\bar{x}_1$. Search direction change, according to the principle mentioned above, does not occur at the partial extremes (denoted by circles in Fig. 6), but at the points $A_1(\bar{x}_3, \bar{\bar{x}}_1)$, $A_2(\bar{x}_3, \bar{\bar{x}}_4)$, $A_3(\bar{x}_5, \bar{\bar{x}}_4)$, $A_4(\bar{x}_5, \bar{\bar{x}}_5)$. Let points A_1, A_2, A_3, A_4 deviate from the partial extreme points by an identical amount equal to $\underline{c}$. Then

$$\begin{aligned} \bar{x}_3 &= \bar{x}_1 + c, & \bar{x}_5 &= \bar{x}_3 + c, \\ \bar{\bar{x}}_4 &= \bar{\bar{x}}_2 + c, & \bar{\bar{x}}_5 &= \bar{\bar{x}}_4 + c. \end{aligned} \tag{13}$$

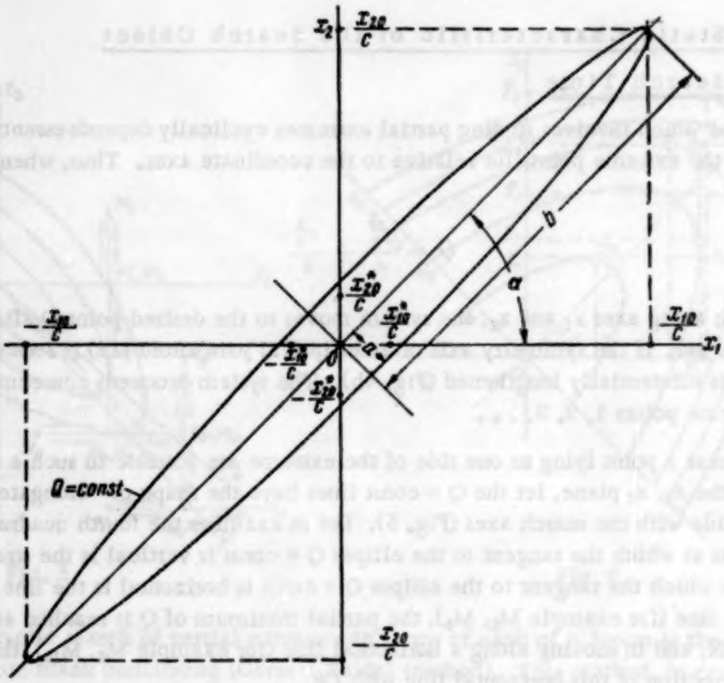


Fig. 7.

In the region for which

$$\Delta x_1 = |\bar{x}_5| - |\bar{x}_3| < 0, \qquad \Delta x_2 = |\bar{\bar{x}}_5| - |\bar{\bar{x}}_1| < 0, \tag{14}$$

the system moves to the extreme point in the search process. When $\Delta x_1 > 0$ and $\Delta x_2 > 0$, the system moves away from the extreme.

The equation of an elliptical paraboloid with its symmetry axes rotated by an angle α is written in the form

$$Q = \frac{(x_1 \cos\alpha - x_2 \sin\alpha)^2}{a} + \frac{(x_1 \sin\alpha + x_2 \cos\alpha)^2}{b} \tag{15}$$

From the equation

$$\frac{\partial Q}{\partial x_1} = 0, \tag{16}$$

where $x_2 = \bar{\bar{x}}_1$, $x_2 = \bar{\bar{x}}_4$, $x_2 = \bar{\bar{x}}_5$, and the equation

$$\frac{\partial Q}{\partial x_2} = 0, \tag{17}$$

where $x_1 = \bar{x}_3$, $x_1 = \bar{x}_5$, we find the coordinates of points A_1, A_2, A_3, and A_4:

$$\bar{x}_3 = \frac{\bar{\bar{x}}_1 (b-a)\,\mathrm{tg}\,\alpha}{b + a\,\mathrm{tg}^2\alpha} + c = A\bar{\bar{x}}_1 + c, \tag{18}$$

where

$$A = \frac{(b-a)\,\mathrm{tg}\,\alpha}{b + a\,\mathrm{tg}^2\alpha},$$

$$\bar{\bar{x}}_4 = \frac{\bar{x}_3 (b-a)\,\mathrm{tg}\,\alpha}{a + b\,\mathrm{tg}^2\alpha} + c = B\bar{x}_3 + c, \tag{19}$$

where

$$B = \frac{(b-a)\,\mathrm{tg}\,\alpha}{a + b\,\mathrm{tg}^2\alpha}.$$

Using (13), (18), and (19), we find

$$\bar{x}_5 = AB\bar{x}_3 + Ac + c, \tag{20}$$

$$\bar{\bar{x}}_5 = A^2B^2\bar{\bar{x}}_1 + AB^2c + ABc + Bc + c. \tag{21}$$

We find the coordinates of the points around which self-oscillations occur (x_{10}, x_{20}) from the equations

$$\Delta x_1 = 0 \text{ and } \Delta x_2 = 0, \tag{22}$$

$$x_{10} = \bar{x}_3 = \frac{c(A+1)}{1-AB}, \tag{23}$$

$$x_{20} = \bar{\bar{x}}_1 = \frac{c(B+1)}{1-AB}. \tag{24}$$

Let $\alpha = \pi/4$, $b/a = 10$. Then $x_{10} = x_{20} = 5.5$ c.

The given system will self-oscillate around the point with the coordinates which have been found.

If c has a negative sign, the system, according to (23) and (24), will self-oscillate around the point with coordinates (Fig. 7)

$$x_{10} = x_{20} = -5.5c.$$

Thus, the point located in searching for the extreme of a two-variable function will be at a considerable distance from the true value. This distance can be greatly reduced by two methods:

1) by changing the optimizer structure so that the sign of the deviation of the system from partial extreme points is changed in a repetitive cycle (by alternating the direction of the first step along each of the variables);

2) by putting a coordinate transformer in the system, which changes the direction of the search axes.

Let us examine the first of these methods. In the method of alternating the first-step direction, equations (13) will take the form

$$\begin{aligned} \bar{x}_3 &= \bar{x}_1 + c, & \bar{x}_5 &= \bar{x}_2 - c, \\ \bar{\bar{x}}_4 &= \bar{\bar{x}}_2 + c, & \bar{\bar{x}}_5 &= \bar{\bar{x}}_3 - c. \end{aligned} \tag{25}$$

Self-oscillations will occur around the point with coordinates

$$x_{10}^{+} = -\frac{c(A+1)}{1+AB}, \tag{26}$$

$$x_{20}^{+} = -\frac{c(B+1)}{1+AB}. \tag{27}$$

As in the case examined earlier, let $\alpha = \pi/4$, $b/a = 10$. Then $x_{10}^{+} = 1.1$ c, $x_{20}^{+} = 1.1$ c.

In this case, the coordinates of the point located will be near the extreme point.

It follows from (26) and (27) that the search error can be greatly reduced by bringing α to zero or, in practice, by diminishing it. This problem will be treated in detail in the next article.

LITERATURE CITED

1. N. E. Kobrinskii, Continuous Action Mathematical Machines [in Russian] (State Press for Technical and Theoretical Literature, 1954).
2. A. A. Fel'dbaum, "On the use of computers in automatic systems," Avtomatika i Telemekhanika 17, No. 11 (1956).
3. R. I. Stakhovskii, "On the comparison of some search methods for an automatic optimizer," Theory and Application of Discrete Automatic Systems [in Russian] (Conference transactions, September 22-26, 1958) (AN SSSR Press, 1960).

4. A. A. Krasovskii, "The dynamics of continuous extremal control systems, based on the gradient method," Izv. AN SSSR, Otd. Tekhn. Nauk No. 3 (1959).
5. A. G. Ivakhnenko, "Automatic control systems with logic elements," Session of the Academy of Sciences, USSR on the Scientific Problems of Industrial Automation, October 15-20, 1956 [in Russian] (AN SSSR Press, 1957).
6. L. N. Fitsner, "On the principles of construction and methods of analyzing certain types of extremal systems," Theory and Application of Discrete Automatic Systems. Conference Transactions, September 22-26, 1958 [in Russian] (AN SSSR Press, 1960).
7. B. White, "The quare optimal controller," Instruments and Automation No. 11 (1956).
8. Ya. Z. Tsypkin, Theory of Pulse Systems [in Russian] (Fizmatgiz, 1958).

ANALOG SOLUTION OF ALGEBRAIC AND TRANSCENDENTAL EQUATIONS BY THE GRADIENT METHOD

M. V. Rybashov
Translated from Avtomatika i Telemekhanika, Vol. 22, No. 1, pp. 77-88, January, 1961
Original article submitted June 2, 1960

This article presents a method for solving sets of finite equations by means of analog computers. An auxiliary set of differential equations is constructed from the initial set of finite equations, and differs from them in that the stable rest points correspond to the roots of the initial set. In this paper, the particulars of this method are studied, and the conditions under which it is most effective are found.

1. Introduction

Modern mathematical indirect analog computers are basically intended for solving various dynamic problems from the class of ordinary differential equations, both by the nature of their general structure and by their individual elements.

This, however, is not the only type of problem which can be successfully solved by these devices. It is known that it is possible, in principle, to solve algebraic and transcendental equations, arising from the solution of certain differential equations under special conditions, by using such computers.

The method of solving algebraic and transcendental equations by such means will henceforth be called the "differential equation" method. Historically, this method arose as a result of work on the elimination of instability which occurred in solving algebraic equations by the implicit function method on continuous action machines.

According to this method, instead of using the initial equation

$$f(x) = 0 \tag{1}$$

one uses either its equivalent form $x = x + f(x)$, or an equation which is nearly like Eq. (1),

$$f(x) = -\frac{x}{k}, \qquad k \gg 1. \tag{2}$$

A block diagram of the problem in the form of Eq. (2) is shown in Fig. 1. It follows from Fig. 1 and Eq. (2) that the structural diagram is a closed-loop network, and that the dynamic properties of the elements are not considered. In view of this, the solution $x^* = \alpha$ (α is a number), which turns Eq. (2) into an identity, may prove to be unstable.

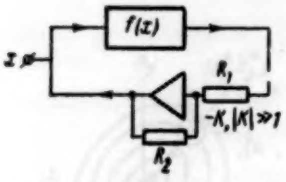


Fig. 1

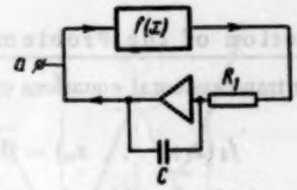


Fig. 2

Thus, the problem arose of how to ensure network stability for sufficiently large k. This problem was complicated by the fact that it is difficult to account for small inertias which reflect the dynamic properties of the network, and in practice, a differential equation was solved which did not correspond to the block diagram of Fig. 1.

However, one remarkable property of dynamic systems has been known for a long time. If one large concentrated inertia is artificially put into such a system with small inertias, the system will be stable in the majority of cases, if the differential equation formed without accounting for the small parameters has a stable solution. The latter statement evidently requires some explanation. In the case being considered, the artificial inertia is created by replacing the high-gain amplifier with an integrator (Fig. 2). In this case, the differential equation, neglecting small parameters, is

$$T\frac{dx}{dt} = -f(x), \tag{3}$$

where $T > 0$ is the integrator constant. Substituting $\tau = t/T$, we get

$$\frac{dx}{d\tau} = f(x). \tag{3'}$$

The hypothesis of the suppression of small inertias can be formulated as follows: If Eq. (3') has a stable solution $x^* = \alpha$, which is the root of Eq. (1), then a number $T \neq 0$ can be found for which the actual system shown in Fig. 2 will be stable at the point $x = x^*$. This definition gives rise to the requirement that the rest point x^* of Eq. (3') be stable. This equation is generally formed without considering this requirement, since its right half is arbitrary.

The hypothesis may be untrue even if the latter condition is met. However, practice shows that if this condition is met, a number T for which the system is stable will exist for a broad class of problems. The method presented for solving algebraic equations is the differential equation method. At the equilibrium point $dx/d\tau = 0$, Eq. (3') turns into Eq. (1), and we get the solution $x = x^*$ at point a (Fig. 2). The derivation presented gives rise also to an inverse relation: the differential equation method is inapplicable in the form described here if the rest point x^* of Eq. (3') is unstable. Practice shows that the majority of equations fall into this class.

A new problem arises: How does one construct a differential equation which will have the same equilibrium points as Eq. (3), and have all these points asymptotically stable?

Methods for constructing such equations are known in the literature [1]. The following equation:

$$\frac{dx}{dt} = -2\lambda f\frac{\partial f}{\partial x}, \qquad \lambda > 0. \tag{4}$$

is written instead of Eq. (3).

The right half of this equation is a derivative of x having the form λf^2, and the equality itself shows that the form is minimized in terms of the gradient. No detailed study of this is given in [1].

This article is devoted to a more detailed study of this method. The peculiarities of the method are explained, the region of its most suitable application is found, and stability in the large is examined. It will be shown that the gradient method, as it applies to the stated problem, is derived from Lyapunov's second method.

2. Statement and Solution of the Problem

Let a set of algebraic or transcendental equations exist:

$$f_i(x_1, \ldots, x_n) = 0 \qquad (i = 1, \ldots, n), \tag{5}$$

where $x_1, x_2, \ldots, x_n$ are the variables for which the given set must be solved. The functions f_i are continuously differentiable over their entire arguments, at least in the bounded region D. In the vicinity of the solution $(\alpha_1, \ldots, \alpha_n)$

$$\det A = \det \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix} \neq 0. \tag{6}$$

The latter condition ensures that the rest point will be isolated. The determinant (6) may equal zero on individual manifolds. These cases, however, will be dealt with specially.

We will use Lyapunov's theorem about the asymptotic stability of the equilibrium point [2] to construct a set of differential equations having the properties we need. According to this theorem, the equilibrium point $(\alpha_1, \ldots, \alpha_n)$ of the set of differential equations

$$\frac{dy_i}{dt} = f_i(\alpha_1 + y_1, \ldots, \alpha_n + y_n) \qquad (i = 1, \ldots, n) \tag{7}$$

is asymptotically stable for sufficiently small values of y_i ($i = 1, \ldots, n$) if a positive-defined function* $V = V(y_1, \ldots, y_n)$ exists whose derivative with respect to t is negative-defined.

The content of the theorem can be written as the inequality

$$\frac{dV}{dt} = \sum_{i=1}^{n} \frac{\partial V}{\partial y_i} f_i < 0. \tag{8}$$

Let us form, by some means, a positive-defined function V of $f_1, \ldots, f_n$:

$$V = V(f_1, \ldots, f_n),$$

where

$$f_i = f_i(\alpha_1 + y_1, \ldots, \alpha_n + y_n) \qquad (i = 1, \ldots, n). \tag{9}$$

We note that it will also be positive-defined in terms of the arguments $y_1, \ldots, y_n$.

The function goes to zero if $f_1 = f_2 = \ldots = f_n$, that is, at the point $(x_1, \ldots, x_n) = (\alpha_1, \ldots, \alpha_n)$, where $x_i = \alpha_i + y_i$ ($i = 1, \ldots, n$), which is a root of the initial set of finite equations (5).

It follows from inequality (8) that, if one forms a set of differential equations of the form

$$\frac{dy_i}{dt} = -\frac{\partial V}{\partial y_i} \qquad (i = 1, \ldots, n), \tag{10}$$

or its equivalent

$$\frac{dx_i}{dt} = -\frac{\partial V}{\partial x_i} \qquad (i = 1, \ldots, n), \tag{10'}$$

the conditions of the theorem will almost be met,** that is, a positive-defined function V exists whose derivative equals

---

* A positive-defined (negative-defined) function V is a function which is positive (negative) everywhere except at the points $y_1 = y_2 = \ldots = y_n = 0$, where $V = 0$.

** Strictly speaking, the conditions of Lyapunov's other theorem — about nonasymptotic stability — will be met.

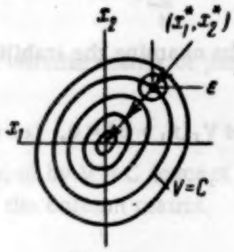


Fig. 3

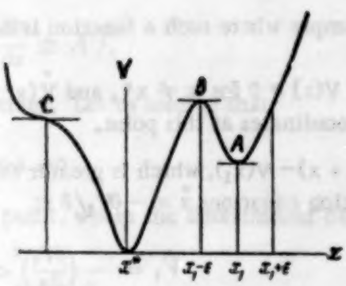


Fig. 4

$$\dot{V} = -\sum_{i}^{n} \left(\frac{\partial V}{\partial y_i} \right)^2,$$

or at least is nonpositive:

$$\dot{V} \leqslant 0.$$

It follows from this that one cannot always guarantee asymptotic stability even in the small. However, if we use the fact that the coordinates of any physical system are always subject to the effect of small random forces, it can be shown that asymptotic stability will exist in the majority of cases.

Indeed, let $dV/dt \equiv 0$ on the manifold M of dimensionality $m < n$, i.e., a dimensionality less than that of the space. Then a positive probability exists that the mapping point will be in space L_1, formed from L by being taken from the manifold M. The inequality

$$\frac{dV}{dt} < 0 \tag{11}$$

holds strictly in the latter space.

If the mapping point, affected by small perturbations, is thrown onto a trajectory moving away from manifold M, Lyapunov's theorem about asymptotic stability will be justified from there on. Let us clarify this fact by an example of a space where n = 2 (Fig. 3). In terms of Eq. (7), let $dV/dt \equiv 0$ on the manifold V = C, and, consequently, the set of equations

$$\dot{x}_1 = -\frac{\partial V}{\partial x_1}, \quad \dot{x}_2 = -\frac{\partial V}{\partial x_2}$$

reduce to the identity $\partial V / \partial x_i \equiv 0$ everywhere on this manifold.

Let us assume that the trajectory intersects V = C at the point (x_1°, x_2°). Because of fluctuations, the mapping point can be in the ϵ-vicinity of this point. But inside $V < C$, $\dot{V} < 0$, and the point, once thrown there, will move away to the equilibrium position.

It is evident that none of these statements is justified if $dV/dt \equiv 0$ on a manifold of dimensionality m = n, or over the whole of space L.

Thus, the use of the hypothesis of small perturbations greatly weakens the conditions of the cited Lyapunov theorem.

Before considering a concrete choice of the function V, it is necessary to examine some possible cases of system behavior for $x_i \in D$, where D, as mentioned earlier, is a bounded region. This examination is obligatory, since it is known [3] that the system behavior near the isolated special point does not fully determine the behavior in a sufficiently large, though bounded, region.

It is evident that the system will be asymptotically stable in the presence of any perturbations $x_{i_0} \in D$, if the structure of V is the same everywhere as in the vicinity of the rest point. Near the rest point, the function V in $\underline{n}$-dimensional space is a one-parameter family of closed surfaces enclosed in one another.

It is easy to give an example where such a function is destroyed if $r = \sum_i |x_{i0}|$ is sufficiently large.

Figure 4 shows a function $V = V(x) > 0$ for $x \neq x^*$, and $\dot{V}(x^*) = 0$. Let us examine the stability of point A. We will locate the origin of the coordinates at this point.

We will form $V_1 = V(x_1 + x) - V(x_1)$, which is greater than zero, and $V_1(x_1 + 0) = 0$. Let us examine its derivative $\dot{V}_1$ in terms of the motion equations $\dot{x} = -\partial V_1/\partial x$;

$$\dot{V}_1 = -\left(\frac{\partial V_1}{\partial x}\right)^2 < 0,$$

where $x \leq |\epsilon|$ and $\dot{V}_1(x_1) = 0$. Asymptotic stability of point x_1 will consequently be observed in the presence of perturbations not exceeding those mentioned. Points B and C are evidently unstable in the presence of fluctuations. Similar arguments can be presented for any finite n. As we see, points A, B, and C are stationary points which are defined by the set of equations $\partial V/\partial x_i = 0$ ($i = 1, \ldots, n$). Consequently, one must attempt to construct the function so that it is geometrically a "smooth cup" in $(n + 1)$-dimensional space, with a minimum at the point under investigation for stability. It is true that this function is ideal. A function can be constructed which has "safe" extremal points like point C (inflection points, saddles, etc.). For this is needed partial derivative $\partial V/\partial x_i \neq 0$ for any x_i, except at points corresponding to the solution of Eq. (5).

The discussions on the construction of the function V presented above can be used in actual cases of equations (5), and are also needed for further derivations.

We note that one can choose a positive-defined* quadratic form

$$V = \sum_{i,j}^{n,n} a_{ij} f_i f_j, \qquad a_{ij} = \text{const}$$

for the function V. The canonical form, in particular, may be used:

$$V = \frac{1}{2}\sum_i^n \lambda_i f_i^2, \qquad \lambda_i > 0 \qquad (i = 1, \ldots, n). \tag{12}$$

It is easy to prove that if $\lambda_i = 1$ ($i = 1, \ldots, n$), the analog set of equations

$$\dot{x}_i = -\frac{\partial V}{\partial x_i} \qquad (i = 1, \ldots, n) \tag{13}$$

can be written in the form

$$\frac{dx}{dt} = -A' f, \tag{14}$$

where dx/dt and f are column matrices, A' is a matrix which is a transponent of the Jacobian matrix A. It follows from Eq. (14) that condition (6) is met at the solution, and all false singularities are on the manifold

$$\det A = 0.$$

False solutions need not be feared if the analog is universal, since they can be isolated. If the solution is false, $\sum_i^n |f_i| \neq 0$. In the case of a specialized analog, when the path of the root, as the left half of (5) changes with time, is important, an "ideal" function V must be constructed by some means or other, or else the root may go off on a false path.

Let us examine some specific properties of Eq. (13). Equating the right halves of (13) and (14), we get

* A positive-defined quadratic form is a quadratic form whose matrix is positive, that is, all the main minors of the matrix are greater than zero.

$$\frac{\partial V}{\partial x} \equiv A'f, \tag{15}$$

where $\partial V/\partial x$ is a column matrix of partial derivatives. Let us assume that

$$\det A \neq 0$$

in the region bounded by V = C (except at the rest point, where the determinant can go to zero). Then it can be demonstrated that the column matrix

$$A'f \neq 0.$$

Indeed, the matrix A' is nondegenerate if only one of the f_i is not zero in the vicinity of the rest point. Consequently, in view of identity (15), only one function $\partial V/\partial x_i$ is nonzero everywhere in this region.

Remembering that

$$\dot{V} = -\sum_i^n \left(\frac{\partial V}{\partial x_i}\right)^2,$$

we get the final result $\dot{V} < 0$, and this means that the conditions of the asymptotic stability theorem are strictly met. The point is asymptotically stable in the presence of perturbations not exceeding the limits of the region being examined.

Let us take an example. Let there be a set of transcendental equations of the form

$$f_1 = x_1 - e^{x_2} + a_1 = 0, \qquad a_1 = \text{const},$$
$$f_2 = e^{x_1} + x_2 + a_2 = 0, \qquad a_2 = \text{const}.$$

It can be shown by graphical construction that this set of equations has a real solution for certain values of a_1 and a_2 (for example, $a_1 = a_2 = 1$). We will explain stability in the small. To do this, we find the roots of equations

$$\Delta(\lambda) = \begin{vmatrix} \frac{\partial f_1}{\partial x_1} - \lambda & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_1}{\partial x_2} & \frac{\partial f_2}{\partial x_2} - \lambda \end{vmatrix} = \begin{vmatrix} 1-\lambda & -e^{x_2} \\ e^{x_1} & 1-\lambda \end{vmatrix} = \lambda^2 - 2\lambda + 1 + e^{x_1+x_2} = 0.$$

$$\lambda_{1,2} = 1 \pm i\sqrt{e^{x_1+x_2}}.$$

The roots have positive real parts, so a set of differential equations of the form

$$\dot{x}_1 = x_1 - e^{x_2} + a_1, \quad \dot{x}_2 = e^{x_1} + x_2 + a_2$$

has an unstable equilibrium point. We will construct a new set of differential equations by using equality (13). We will use $V = \frac{1}{2}(f_1^2 + f_2^2)$ as the function V. Then

$$\dot{x}_1 = -\left(f_1\frac{\partial f_1}{\partial x_1} + f_2\frac{\partial f_2}{\partial x_1}\right), \quad \dot{x}_2 = -\left(f_1\frac{\partial f_1}{\partial x_2} + f_2\frac{\partial f_2}{\partial x_2}\right).$$

Substituting the expressions for f_i, we finally get

$$\dot{x}_1 = -[(x_1 - e^{x_2} + a_1) + (e^{x_1} + x_2 + a_2)e^{x_1}],$$
$$\dot{x}_2 = -[-(x_1 - e^{x_2} + a_1)e^{x_2} + (e^{x_1} + x_2 + a_2)].$$

In the example being considered, $\det A \neq 0$ for any values of (x_1, x_2). Indeed,

$$\det A = 1 + e^{x_1+x_2} \neq 0$$

and, consequently, the rest point, being a solution of the transcendental set, is asymptotically stable in the large,* that is, for any initial conditions.

Equation (14) comes as a result of Lyapunov's theorem. This theorem gives the sufficient conditions for the existence of stability, and thus it is desirable to try to simplify this set of equations. Practice shows that one of the simplifications is to make all the elements of matrix A zero except along diagonals; this often gives positive results. The analog system in this case assumes the form $\dot{x}_i = -\partial f_i / \partial x_i$ $(i = 1, \ldots, n)$. This is significantly simpler than (13).

3. Finding the Roots of Analytic Functions by the Gradient Method

The statements presented in the preceding section on the properties of an "ideal" function V permit one to state the following problem. Let a function V of f_i $(i = 1, \ldots, n)$ be given by formula (12) $(\lambda_i = 1;\ i = 1, \ldots, n)$. It is necessary to find a class of function which, when substituted in this formula, will give an "ideal" function V. It happens that one of these classes is that consisting of the real and imaginary parts of any analytic function of one complex variable. Let a real analytic function $f(p)$ of complex argument $p = \alpha + i\beta$ be given. The problem is to find its roots $\alpha_k + i\beta_k (k = 1, \ldots, q)$ by the gradient method. To do this, we form the function

$$V = \frac{1}{2}\left[M^2(\alpha, \beta) + N^2(\alpha, \beta)\right], \tag{16}$$

where $M(\alpha, \beta)$ and $N(\alpha, \beta)$ are functions corresponding to the representation of the initial function $f(p)$ in the form

$$f(\alpha + i\beta) = M(\alpha, \beta) + iN(\alpha, \beta).$$

The function f goes to zero, corresponding to the root $\alpha_k + i\beta_k$ when, and only when

$$M(\alpha, \beta) = 0, \quad N(\alpha, \beta) = 0. \tag{17}$$

Let us solve (17) (which is generally nonlinear) by the gradient method. We will first minimize (16). It is evident that it will be zero when, and only when, equality (17) is satisfied. Let us examine the set of differential equations

$$-\frac{d\alpha}{dt} = M\frac{\partial M}{\partial \alpha} + N\frac{\partial N}{\partial \alpha}, \quad -\frac{d\beta}{dt} = M\frac{\partial M}{\partial \beta} + N\frac{\partial N}{\partial \beta}, \tag{18}$$

which are formed from (10).

It follows from Liouville's theorem [4] that V is an infinitely large function ($V \to \infty$ for $|\alpha| + |\beta| \to \infty$), since every analytic function possesses this property, providing it is not a constant and is analytic everywhere.

This fact permits one to state that, for any initial conditions, the mapping point goes to one of the minimums of V during a course of time, and goes to one of the roots of this function if there are no false minimums; that is, minimums which do not correspond to roots of $f(p)$.

If there are no false roots, V is shaped like a series of one-sided cups with minimums at roots, interconnected by saddles. The entire analytic region in the (α, β) plane is divided into a series of contiguous subspaces connected at their roots. Let us proceed with the proof itself. Suppose that the function V has no minimums or maximums at points defined by the inequality $f(p) \neq 0$, and that a two-dimensional continuum of stationary points of the other type does not exist. We note that the function V is one-half the square of the modulus of $f(p)$. It follows from the principle of the maximum modulus of a function of a complex variable that V has no maximums or minimums at any point in the analytic region except at roots, where $|f(p)| = 0$. We will now show that the analytic function does not have a continuum of points at which

$$\frac{\partial V}{\partial \alpha} = \frac{\partial V}{\partial \beta} \equiv 0. \tag{19}$$

---

*In judging the stability in the large, one can see that the function V is "infinitely large" [3]. A sufficient condition of an infinitely large function is that $V \to \infty$ for $\sum_i^n |x_i| \to \infty$. This condition is met in the example presented.

Let condition (19) be met in some region D. It can then be proved that V is constant on this continuum (see Appendix). It is known [4] that, if the modulus of an analytic function is constant somewhere in a region G belonging to the analytic region, then V = const over the entire analytic region. We have a function $V \neq$ const. In this case, the derivative $\frac{dV}{dt} = -\left[\left(\frac{\partial V}{\partial \alpha}\right)^2 + \left(\frac{\partial V}{\partial \beta}\right)^2\right]$ goes to zero on a manifold of dimensionality less than that of the space, which is what we wanted to prove. Thus, it has been established that the solution will converge to one of the roots of the function $f(\alpha + i\beta)$ if there are no perturbations in α and β exceeding the limits of the analytic region. It follows from this that, in using the gradient method, the problem of finding the roots of analytic functions reduces to one of finding all contiguous regions; these completely fill the (α, β) plane.

As is known, integral functions, polynomials in particular, are related to analytic functions. Consequently, the well-known problem of finding the roots of a polynomial can be solved by the gradient method.

The case of finding the complex roots of one equation has been examined above. It would appear that a set of differential equations will possess the property of not having false rest points even in the multidimensional case, if we take $V = \sum_i^n [M_i^2(\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n) + N_i^2(\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n)]$ as the Lyapunov function. Here, M_i and N_i are the real and imaginary parts of the ith equation of the set $f_i(p_1, \ldots, p_n) = 0$, respectively. It is easy to prove, however, that such a function is not the modulus of an analytic function of many complex variables, and thus one cannot guarantee that the set

$$\frac{d\alpha_i}{dt} = -\frac{\partial V}{\partial \alpha_i}, \quad \frac{\partial \beta_i}{\partial t} = -\frac{\partial V}{\partial \beta_i} \qquad (i = 1, \ldots, n)$$

will not have false rest points. We will mention here one method which may give positive results. Let us consider the set of differential equations

$$\frac{d\alpha_i}{dt} = -\frac{\partial V_i}{\partial \alpha_i}, \qquad V_i = M_i + N_i,$$
$$\frac{d\beta_i}{dt} = -\frac{\partial V_i}{\partial \beta_i} \qquad (i = 1, \ldots, n).$$

For any ith subsystem of equations, the variables α_k, β_k $(k \neq i)$ are independent parameters, which deform the function V in time. The movement of the mapping point in the plane of subspace (α_i, β_i) is a function of the gradient. However, if the rate of deformation is significant, the movement may cause the rank of $V = C_j$ to increase, i.e., instability will be observed.

4. The Effect of Bounding the Linearity Limits of the Analog Elements on the Stability of the Whole System

Let us examine the block diagram of a network for solving a set of two equations* (Fig. 5). It consists of two functional converters $\Phi_1 = -\frac{\partial V}{\partial x_1}$ and $\Phi_2 = -\frac{\partial V}{\partial x_2}$ and two integrators. Two loops I and II are clearly shown in the network, corresponding to the differential equations $\dot{x}_1 = \Phi_1$ and $\dot{x}_2 = \Phi_2$. Initial conditions are introduced in the usual way — capacitances C_1 and C_2 are charged to x_{10} and x_{20} (with switches K_1 and K_2 open). After closing K_1 and K_2, two modes must be distinguished:

1) integrators I_1 and I_2 operate in the linear region when $t > 0$;

2) the integrators operate outside the linear region part of the time.

The second mode is related to cases when the absolute magnitudes of the coordinates $|x_i|$, for $t > 0$, at certain times assume values greater than the initial ones. The functional converters may either operate linearly, or may depart from linearity, in both cases.

*If $n > 2$, the network remains similar, except that the number of closed loops increases in proportion to n.

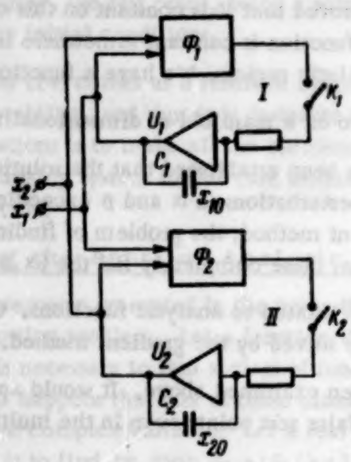


Fig. 5.

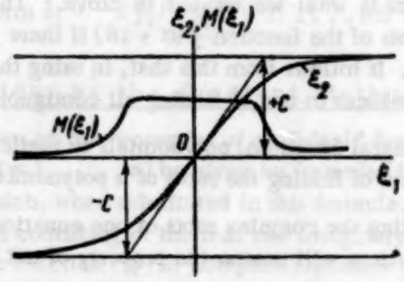


Fig. 6.

Let us investigate both cases from the standpoint of the stability of the rest point defined by the equations $\Phi_1 = 0$, $\Phi_2 = 0$.

We will examine the first mode. Let the linear zones of the functional converters be max $|\Phi_i| = A_i$, $A_i > 0$ $(i = 1,2)$. It is also known that if $|\Phi_1| < A_1$ and $|\Phi_2| < A_2$, the process converges to the roots (x_1^*, x_2^*) of the equations $\Phi_1 = 0$, $\Phi_2 = 0$. We are only interested in the behavior of the system in the nonlinear region. Let us examine the function*

$$V = \frac{1}{2}(f_1^2 + f_2^2),$$

where f_1 and f_2 are the initial functions of the unconverted system $f_1 = 0$, $f_2 = 0$, and its time derivative is

$$\frac{dV}{dt} = \frac{\partial V}{\partial x_1}\dot{x}_1 + \frac{\partial V}{\partial x_2}\dot{x}_2.$$

But $\dot{x}_i \equiv A_i \operatorname{sign} \frac{\partial V}{\partial x_i}$ in the nonlinear zone $\left(\operatorname{sign} \frac{\partial V}{\partial x_i} \equiv 1 \text{ when } \frac{\partial V}{\partial x_i} > 0 \text{ and sign } \frac{\partial V}{\partial x_i} \equiv -1 \text{ when } \frac{\partial V}{\partial x_i} < 0\right)$. Taking this into account, we get

$$\dot{V} = -\left(\frac{\partial V}{\partial x_1} A_1 \operatorname{sign} \frac{\partial V}{\partial x_1} + \frac{\partial V}{\partial x_2} A_2 \operatorname{sign} \frac{\partial V}{\partial x_2}\right) < 0.$$

Thus, the first mode is safe, and convergence of the process to the roots of the initial system is ensured. We will examine the second mode, assuming that the nonlinear functions $\xi_2(\xi_1)$ are smooth and approach the saturation boundary $|C|$ asymptotically (Fig. 6). We will write ξ_2 in the form

$$\xi_2 = \frac{\xi_2(\xi_1)}{\xi_1}\xi_1 = M(\xi_1)\xi_1.$$

The function $M(\xi_1)$ is positive everywhere, and approaches the axis of ξ_1 asymptotically as $|\xi_1| \to \infty$. We will assume that a block with a transfer function $M(\Phi_i)$ is connected in series with the functional converter Φ_i. Let us write the differential equation of the integrator, taking into account the nonlinearity of the differential coefficient $\underline{k}$. As is known, the transfer function of the integrator has the form

$$W(p) = -\frac{k}{Tkp+1}, \quad T > 0, \quad k > 0, \quad p \equiv \frac{d}{dt}.$$

We will assume that the amplification factor $\underline{k}$ is a function of the output coordinate $\underline{x}$, in the same manner as $M(\xi_1)$. Consequently, the integrator equation is

$$T\frac{dx}{dt} + \frac{x}{k(x)} = -y,$$

*It is assumed that V has the "ideal" form.

where $\underline{y}$ is the input and T is the integrator constant. The analog set of equations is then written in the form

$$\frac{dx_i}{dt} + \frac{x_i}{k(x_i)} = -M\left(\frac{\partial V}{\partial x_i}\right)\frac{\partial V}{\partial x_i} \quad (i = 1, 2), \quad T = 1. \tag{20}$$

Considering (20), the derivative of V with respect to $\underline{t}$ can be variable in its sign:

$$\dot{V} = -\sum_{i=1}^{2}\left[M\left(\frac{\partial V}{\partial x_i}\right)^2 - \frac{\partial V}{\partial x_i}\frac{x_i}{k(x_i)}\right].$$

For large x_i, i.e., for significant departure from the linear zone, the second term may be large, and may determine the sign of the derivative $\dot{V}$. This kind of mode is dangerous – false rest points may appear (for example in the saturated state), or self-oscillations may arise. When this mode appears, the scale of x_i must be altered in such a way as to decrease the range of change.

One can draw the following conclusions on the basis of the material presented:

1. In solving a set of finite equations with an analog computer, it is desirable to construct the initial differential equations by means of the gradient method. This gives a logical approach to the composition of the differential equations, and ensures convergence to the rest point, which is the solution of the set of finite equations, in the majority of cases.

2. The rest points of the set of differential equations must be asymptotically stable. It has been shown in this paper that, due to the positive effect of small perturbations, the gradient method ensures asymptotically stable rest points in the majority of cases.

3. A sufficient condition of asymptotic stability of the equilibrium point when V is expressed in the form $V = \frac{1}{2}\sum_{k}^{n} f_k^2$, and when the perturbations do not exceed the limits of a contiguous region, is that the determinant of the Jacobian matrix be nonzero throughout this region.

4. Asymptotically stable rest points of the set of differential equations may include some which are not roots of the initial set of finite equations.

5. It has been proven in this paper that, if the left half of a given finite equation is an analytic function, a set of equations whose independent variable may be complex will not have asymptotically stable false rest points.

6. If integrators and functional converters operate nonlinearly, two operating modes must be distinguished: a safe mode, in which only the functional converters are permitted to operate nonlinearly, and a dangerous mode in which the integrators operate nonlinearly. Stability may be lost in the latter case.

APPENDIX

If a differentiable function V(x,y) meets the conditions $\partial V/\partial x = 0$, $\partial V/\partial y = 0$ at all points in a simply connected region D, then V = const in this region. Let us examine two different points (x,y), (x',y'), and $\Delta V = V(x',y') - V(x,y)$. The points (x,y), (x',y') are so chosen that the triangle (x,y), (x,y'), (x',y') lies entirely in D. Let us assume the contrary, that is, that $V \neq$ const. Then $\Delta V \neq 0$. We will form the increment ΔV in the following way:

$$\Delta V = [V(x', y') - V(x, y')] + [V(x, y') - V(x, y)].$$

Using Lagrange's theorem, we get

$$\Delta V = \frac{\partial V(y', \xi)}{\partial x}(x' - x) + \frac{\partial V(\eta, x)}{\partial y}(y' - y), \tag{21}$$

where $x < \xi < x'$ and $y < \eta < y'$. The points (ξ, y'), $(x, \eta) \in D$ and, consequently,

$$\frac{\partial V(y', \xi)}{\partial x} = \frac{\partial V(\eta, x)}{\partial y} = 0.$$

It follows from (21) that $(x' - x) = (y' - y) = 0$, which contradicts the initial assumption that two different points were chosen. Thus, the theorem is proved by reductio ad absurdum.

LITERATURE CITED

1. F. Murray, Theory of Mathematical Computers [Russian translation] (IL, 1949).
2. A. M. Lyapunov, Collected Works [in Russian] (AN SSSR Press, 1956) Vol. 2.
3. N. N. Krasovskii, Certain Problems Concerned with the Theory of Stable Motion [in Russian] (Fizmatgiz, 1959).
4. I. I. Privalov, Introduction to the Theory of Functions of a Complex Variable [in Russian] (State Press for Technical and Theoretical Literature, 1954).

A METHOD FOR SELECTING THE OPTIMUM STRUCTURE OF A DIGITAL ANALOG COMPUTER

A. V. Shileiko

Translated from Avtomatika i Telemekhanika, Vol. 22, No. 1, pp. 89-96, January, 1961

Original article submitted May 21, 1960

This article presents a method for choosing the optimum structure of digital analog computers on the basis of a special factor – the figure of merit. Comparable structures are described by a series of numbers called the structure parameters. The problem is to find some group of parameters which give the extreme value of the figure of merit. A concrete example of a design of a digital analog computer is given.

Specialized computers intended for operation in automatic control systems must satisfy a number of specific requirements on accuracy, speed of response, reliability, size, operating convenience, etc. To a great degree, satisfying these requirements depends on the rational choice of a structure for the computer. Therefore, a comparative evaluation of computer structures and the problem of an optimum structure in one or another manner are of direct interest in rationally constructing control apparatus.

A special class of computers, called "digital analog" computers, deserves attention in connection with the solution of control problems [1,2]. Digital analog computers are distinguished by a number of advantages in apparatus and use. On the other hand, the majority of computers of this type described in the literature, which are the series type, have a slow response. Attempts to raise the response speed by using parallel and series-parallel structures lead to a rapid growth in the number of elements which, in turn, makes the apparatus more expensive and requires additional measures to ensure reliability. These contradictions make the problem of choosing the optimum structure of digital analog computers especially vexing.

This article is devoted to an examination of this problem. However, it does not pretend to be a complete investigation of the subject, and is only one of the first steps in this direction.

Statement of the Problem

Speed of response, accuracy, and reliability of computers depend primarily on the characteristics of the elements used in them. However, in the majority of cases, the computer designer is limited by the elements readily available, chosen on an economic basis, or representing the best made available by the state of the art. Under such conditions, the required characteristics of the computer may be obtained only by using a number of elements, and by suitably interconnecting them, that is, by changing the computer structure.

Henceforth, we will mean a system composed of standard elements working as a single computer when we speak of the structure of a computer. Standard elements, in turn, mean those which are comparable in complexity,

construction, and function; these are monostable multivibrator summators, memory cells with a one-bit capacity, logic control elements, etc.

By varying the number of elements in the structure, it is possible to get a relative enhancement of one computer characteristic at the cost of worsening others. Thus, for example, increasing the number of parallel-acting computer channels* ordinarily increases the response speed of accuracy, but simultaneously lowers the reliability and makes the computer more difficult to use. Therefore, it is natural to compare different structures on the basis of an over-all examination of the characteristics which are basic to the given type of computer, thereby determining the relative efficiency of the structures. If the structure efficiency is evaluated in some special way by a quantity — the figure of merit — equivalent structures will have the same figure of merit.

On the basis of these considerations, the problem of choosing the optimum structure of a digital analog computer is examined below. By optimum structure, we mean one which satisfies the given technical requirements, and which is characterized by an extreme value of the figure of merit.

In the first part of this paper we present a choice, and the basis for it, of the figure of merit for a structure. In the second part, a method for choosing the optimum structure, using the figure of merit which will have been derived, is given. The third part deals with an examination of a concrete example of choosing the optimum structure. Henceforth, it will be assumed that an algorithm for the structure is given.

Figure of Merit

The basic task of a computer in an automatic control system is to process information fed to it from certain parts of the system. The results are transmitted to other parts of the system and are used in the over-all control process. From this standpoint, the computer can be considered as a communication channel, and its quality or efficiency of structure can be evaluated as the traffic capacity of such a channel. Of course, we consider that the traffic capacity is, to some degree, an idealized characteristic, and that the real amount of information transmitted by the channel per unit time depends on a number of factors, primarily the type of problem solved by the machine, the computer program chosen, etc. These factors, in turn, depend on the way the computer is used, as well as how rigidly the computer links the input and output signals, independently of their information content.

The traffic capacity increases as the number of parallel computer channels is increased and, consequently, cannot fully characterize the quality of the computer. Considering this, computer quality or efficiency should be evaluated by means of its traffic capacity referred to some quantity which characterizes the amount of apparatus used to get such a capacity or, in other words, by the apparatus complexity of the computer.

The better of two structures having the same traffic capacity will be the one which has the fewer number of elements, other conditions being equal. However, different types of standard elements are not completely equivalent from the standpoint of reliability, cost, convenience of use, etc. Hence, the apparatus complexity of a computer is more correctly evaluated by an expression of the form

$$\lambda = a_1 n_1 + a_2 n_2 + \ldots + a_k n_k, \tag{1}$$

where n_i is the number of standard elements of the $\underline{i}$th type used in the computer, and a_i is a coefficient selected on the basis of experience with analogous computers, or other available data.

The a_i coefficients can be made proportional to the probabilities of failure of each of the standard element types. If one then assumes that redundancy is not introduced into the structure to increase reliability, the quantity λ will be proportional to computer failure and, consequently, will be inversely proportional to its reliability. In the literature on reliability [3,4], the λ computed under similar conditions has been called the complexity coefficient, and relation (1) has been called the complexity relation.

The material considered permits one to express the figure of merit of a computer in the form

$$\eta = \frac{c}{\lambda}, \tag{2}$$

where $\underline{c}$ is the traffic capacity of a computer which has a given structure.

*We mean here an increase in the number of parallel channels not intended to introduce redundance.

If, in choosing the figure of merit, not only the traffic capacity of a structure, but its reliability is considered as well, and the possibility of raising reliability by introducing redundancy is taken into account, the figure of merit will certainly be different. The situation becomes more complicated if the cost of building and operating the machine is considered along with reliability. However, it can be assumed that both reliability and cost, if given types of standard elements are used, will be determined basically by the number of these elements, and expression (2) can be taken as the simplest of a number of analogous evaluations.

Method for Choosing the Optimum Structure

The method given in this article for choosing the optimum structure is as follows. After the figure of merit is gotten, some initial structure is composed which satisfies the given algorithm and is used as a first approximation. A number of quantities called structure parameters are introduced on the basis of an analysis of this structure. One must keep in mind that the parameters chosen must describe the structure accurately enough so that it may be distinguished from other possible structures which satisfy the given algorithm. That this may be done will be shown below in a concrete example.

After composing the initial structure and choosing its parameters, the basic functional dependencies relating the figure of merit to the structure parameters are derived. The possible changes which can be made in each of these parameters are determined by analyzing the technical requirements on the computer. Then the problem of choosing the optimum structure can be formulated in the following manner: find the maximum of some function

$$\eta = G(\overline{F}) \tag{3}$$

under the additional conditions

$$\begin{aligned} g_1(\overline{F}) &\geqslant m_1, \\ g_2(\overline{F}) &\leqslant m_2, \\ g_3(\overline{F}) &= m_3, \\ &\ldots\ldots, \end{aligned} \tag{4}$$

where η is the figure of merit, $F = \|F_1, \ldots, F_i, \ldots\|$ are the structure parameters, and m_i are the boundary conditions.

Such a problem may be solved by one of the existing methods, such as the gradient method [5]. In connection with this, it should be mentioned that conditions (4) determine the region of change permitted in the structure parameters, and functions (3) can reach a maximum both inside this region and on its boundary. Under certain conditions, the occurrence of a maximum on the boundary is sufficient cause for reexamining the initial assumptions, or for altering the technical requirements.

The group of parameters obtained is adjusted on the basis of a number of additional considerations which either could not be expressed in (4), or were neglected in the basic solution in order to simplify the problem.

An Example of Optimum Structure Choice

The method described was used for planning a digital analog analyzer consisting of M independent blocks — integrators — which performed the trapezoidal integration approximation

$$z = \sum_{i=1}^{p} \left(y_{i-1} + \frac{1}{2}\Delta y_i \right) \Delta x_i, \tag{5}$$

where $\underline{z}$ is the approximate value of $\int_{x_0}^{x_p} y(x)\,dx$, y_i is the value of $y(x)$ at the end of the $\underline{i}$th division of the $\underline{x}$ axis, Δy_i is the increment of $\underline{y}$ at the $\underline{i}$th division of the $\underline{x}$ axis, and Δx_i is the size of this division. After computer numbers [1] have been introduced, expression (5) is converted to a set of recurrent functions:

$$Y_i^* = Y_{i-1} + \frac{1}{2} a_y(t_i),$$

$$R_i + A a_z(t_i) = R_{i-1} + Y_i^* a_x(t_i),$$

$$a_z(t_i) = \begin{cases} 1 \text{ for } & R_{i-1} + Y_i^* a_x(t_i) \geqslant A, \\ 0 \text{ for } & 0 < R_{i-1} + Y_i^* a_x(t_i) < A, \\ -1 \text{ for } & R_{i-1} + Y_i^* a_x(t_i) < 0, \end{cases}$$

$$Y_i = Y_i^* + \frac{1}{2} a_y(t_i) \tag{6}$$

(in this case, A > R > 0), where Y is a number stored in the Y registers of the integrators; Y* is a number transmitted to the R register of an integrator; R is a number stored in the R register; $a_x(t_i)$ is an input signal representing increment of x; $a_y(t_i)$ is an input signal representing a y increment; $a_z(t_i)$ is the integrator output, representing a z increment, and A is the maximum of Y and R (see [2]).

The outputs of each integrator occur as delta-modulation pulses with three possible values (−1, 0, 1). The expression for the traffic capacity of such a channel was derived by V. M. Baikovskii. It is

$$c = \frac{1}{\delta_t} \log_2 \left(1 + 2\cos\frac{\pi}{A+1}\right) \text{ bits/sec}, \tag{7}$$

where A is the number of quantum steps in the signal, which equals, in this case,

$$A = 2^{N-2}, \tag{8}$$

where N is the total number of stores in the y and R registers (see below) and δ_t is the time between two successive outputs or the duration of one iteration.

If the number A is sufficiently large, one can write, approximately,

$$c \approx \frac{1}{\delta_t} \log_2 3 \text{ bits/sec} \tag{9}$$

and the final form of the expression for the figure of merit of a digital analog analyzer consisting of M independent integrators is

$$\eta = \frac{Mc}{\lambda} = \frac{M \log_2 3}{\lambda \delta_t}. \tag{10}$$

A block diagram of an initial structure which satisfies algorithm (6) is shown in Fig. 1. The numbers Y and R, increments $a_z(t_i)$, and the program which determines the order of interconnection of the individual blocks are stored in memory block MB. The operations required by formulas (6) are performed in adder AD, namely, 3M successive additions at each step. The control apparatus C connects the blocks and sets the scale of the respective variables.

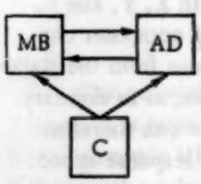


Fig. 1.

The operations of computing formulas (6) are always done in the same sequence. Therefore, address to the memory will always occur in a definite sequence. This eliminates the need for an address system, and makes a delay line or shift register the best memory to use. The structure of the memory, under these conditions, can be described by the following set of parameters: the capacity F_1 of the part of the memory which stores Y and R, and the capacity F_2 of the part of the memory which stores the increments $a_z(t_i)$, the program, and the scale factors.

In the simplest case, adder AD may be a single monostable multivibrator binary summator and a delay line which stores the transfer signal. In the general case, the adder will consist of a number of summators connected in parallel. It is also possible to have the three additions of (6) done by three independent summators connected in series. Starting from this basis, the adder structure may be described by two parameters: the number of parallel channels F_3, and the number of summators connected in series in each channel, F_4. The parameter F_3 is also related to the memory, since the number of parallel channels in the memory must correspond to the number of parallel channels in adder AD.*

* The case may certainly arise where AD and MB have a different number of parallel channels, and appropriate converting elements are connected between them. However, this variation is clearly not rational for such a simple machine as a digital analog computer, and will not be considered here.

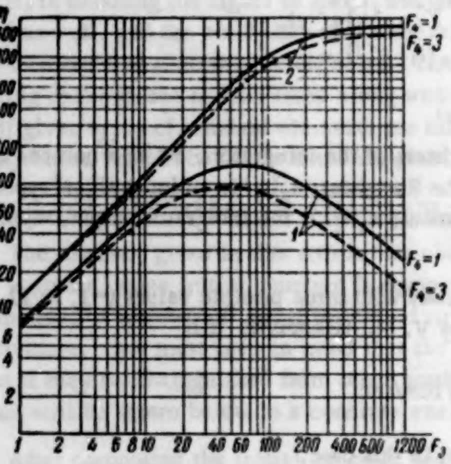


Fig. 2.

The control apparatus of the digital analog computer is simple, and can be physically realized in a number of ways. In terms of the given analysis, it can be characterized by a single structure parameter, namely, the number of elements used, F_8. We also need to know the values of the a_i coefficients in the expression for the complexity coefficient (1), and also data on the speed of response, so that the time for one iteration can be found; i.e., in essence, the characteristics of the elements must be introduced. At present, such complete data on all the different types of elements do not exist. Therefore, we will only consider certain basic types of elements here, and rough estimates of their characteristics will be used.

Magnetic drums, delay lines (magnetostrictive or electromagnetic), diode-ferrite shift registers, ferrite core matrix memories, and transistor-capacitor memory cells will be considered as possible memory elements. Only transistor-diode elements will be considered for logic circuit elements, in particular the control apparatus and adder. The unit for the number of elements is assumed to be an active element consisting of a logic cell which converts or remembers one bit of information for one cycle. Appropriate factors are introduced when other elements are examined. Thus, for example, 10 active elements per single information storage channel are assumed for a magnetic drum.

The response speed of the elements is evaluated in terms of the maximum permissible cyclic frequency for the memory elements, and the delay introduced for the logic elements. Thus, the maximum cyclic frequency is set at 100 kc for magnetic drums, magnetostrictive delay lines, diode-ferrite shift registers, and matrix memories. In the case of electromagnetic delay lines, the maximum cyclic frequency is determined by the auxiliary electronic apparatus, and is set at 1-3 Mc. The maximum cyclic frequency is 1 Mc for transistor-diode memories. The delay introduced by a transistor-diode logic element is estimated to be 0.1 μsec.

In finding the dependence of the figure of merit on the structure parameters, we also take into account the fact that when $F_3 = M$ and $F_3 = NM$, where $N = 2 + \log_2 A$ – the full number of binary steps in X, Y, and R, including sign and additional minor rounding-off operations – the structure of the digital analog computer can be completely represented as an assemblage of individual functional blocks which are autonomous from the standpoint of physical realization. In this case, the individual blocks can be interconnected by busses, as in ordinary computer practice, and scales can be set manually by means of switches or similar devices. We can therefore assume that when F_3 has the values mentioned, F_2 and F_8 go to zero. Finally, in evaluating the response speed, we take into account the fact that, if the number of parallel additions does not exceed N, the summators operating in parallel must be in a cyclic transfer loop.

Figure 2 shows the figure of merit as a function of F_3 and F_4 for the cases when a magnetic drum or transistor-capacitor shift register is used as the memory. In calculating, the most frequently encountered values, M = = 60 and N = 20, were used. It can be seen from the graph that a pure parallel structure is better when shift registers are used, while a series-parallel structure with the number of channels equalling the number of integrators is preferable if a magnetic drum is used. The series-parallel structure is especially convenient if the response speed is not critical.

A block diagram of a digital integrator, designed by means of the method described, is shown in Fig. 3. The integrator consists of N total binary summators Σ_1-Σ_N, in a cyclic transfer loop, and N shift registers P_1-P_N with a capacity of two bits. One step of the computations in (6) is performed in eight cycles corresponding to the times when the shift pulses are fed to registers P. In the first cycle, the operation

$$Y_i^* = Y_{i-1} + \frac{1}{2} a_y(t_i)$$

is performed, the number Y_{i-1} is fed from the registers to the summator inputs A, and the number $a_y(t_i)$ is shifted one step to the right, and is fed into the B input. Information is transmitted inside the registers in the second cycle. In the third cycle, the operation

$$R_i + Aa_z(t_i) = R_{i-1} + Y_i^* a_x(t_i)$$

is performed.

The number R_{i-1} is fed from the registers to the summator inputs A. The number Y_i^* enters the summator inputs B through gates K_1-K_N if $a_x(t_i) = 1$, and through inverters I_1-I_N and gates K_1'-K_N' if $a_x(t_i) = -1$. If $a_x(t_i) = 0$, gates K and K' are closed, and R_{i-1} is added to zero. The resulting signal $a_z(t_i)$ goes to the logic block. In the fourth cycle, information is transmitted inside the registers. In the fifth cycle, the operation

$$Y_i = Y_i^* + \frac{1}{2} a_y(t_i)$$

is performed.

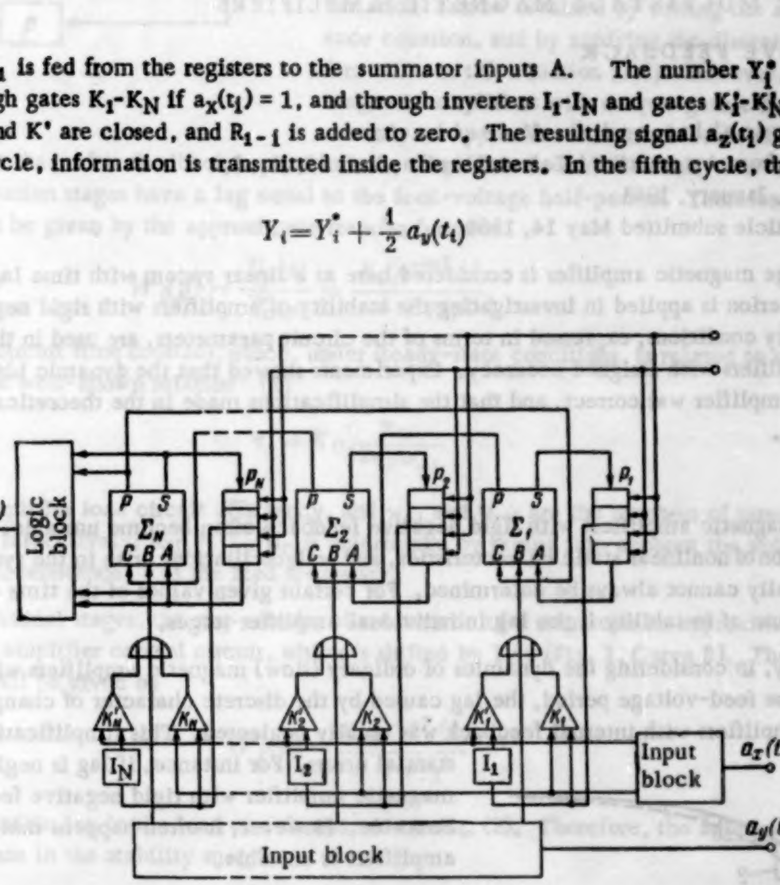


Fig. 3.

The network operates in the same manner as in the first cycle. In the sixth, seventh, and eighth cycles, information is transmitted inside the registers. Moreover, $a_z(t_i) = +1$ signals are emitted from the logic block in the sixth cycle (if such a signal was formed during the given iteration), and $a_z(t_i) = -1$ signals are emitted during the seventh cycle. The numbers Y and R are again fed to the registers from outputs S of the summators after the operations have been performed. Operations with negative numbers are done with an inverse code. The signals $a_x(t_i)$ and $a_y(t_i)$ which enter the integrator during one iteration are stored in the input block. The structure parameters F_3 = NM and F_4 = 1 correspond to a digital analog computer consisting of such integrators.

SUMMARY

This paper describes a method for choosing the optimum structure of a specialized computer based on the derivation of a figure of merit. In one or another way, the choice of this figure is always based on intuitive considerations, and so the method presented cannot give the best of all possible systems in the strict sense of the word. However, it gives an approximation to the best structure, the degree of which depends on a number of factors, and to a large degree on experience in designing such systems.

LITERATURE CITED

1. Digital Differential Analyzers (ed. by B. Ya. Kogan) [Russian translation] (IL, 1959).
2. A. V. Shileiko, "Digital analog computers — a review," Avtomatika i Telemekhanika 20, No. 12 (1959).
3. Electronic Apparatus Reliability Problems (Appendix), Collection [Russian translation] (Oborongiz, 1960).
4. G. G. Friddell and G. G. Jacks, "Operating efficiency of systems (reliability, quality, operating convenience)," Collection: Electronic Apparatus Reliability Problems [Russian translation] (Oborongiz, 1960).
5. R. R. Brown, "A generalized computer procedure for the design of optimum communication systems," Communication and Electronics, Part I, 11, No. 43, 44 (1959).

STABILITY OF MULTISTAGE MAGNETIC AMPLIFIERS WITH NEGATIVE FEEDBACK

M. A. Rozenblat and G. V. Subbotina

Translated from Avtomatika i Telemekhanika, Vol. 22, No. 1,
pp. 97-106, January, 1961
Original article submitted May 14, 1960

A multistage magnetic amplifier is considered here as a linear system with time lag. The frequency criterion is applied in investigating the stability of amplifiers with rigid negative feedback. The stability conditions, expressed in terms of the circuit parameters, are used in the design of decision amplifiers with assigned accuracy. Experiments showed that the dynamic idealization of the magnetic amplifier was correct, and that the simplifications made in the theoretical analysis were permissible.

Multistage magnetic amplifiers with rigid negative feedback often become unstable. In this, the system passes into the region of nonlinear static characteristics, and self-oscillations arise in the system. The main cause of instability naturally cannot always be determined. For certain given values of the time constant and the feedback factor, the cause of instability is the lag in individual amplifier stages.

Until recently, in considering the dynamics of ordinary (slow) magnetic amplifiers with time constants much larger than the feed-voltage period, the lag caused by the discrete character of changes in the output voltage of magnetic amplifiers with internal feedback was usually neglected. This simplification can lead to substantial errors. For instance, if lag is neglected, a two-stage magnetic amplifier with rigid negative feedback must always be stable. However, it often happens that the operation of such amplifiers is unstable.

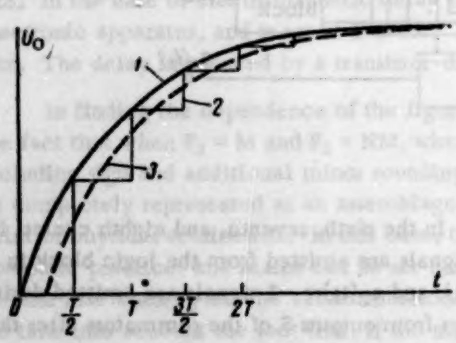


Fig. 1.

The present paper is concerned with the determination of stability conditions and the clarification of related problems in the design of multistage amplifiers which have magnetic amplification stages with internal feedback. It is assumed that each stage operates on the linear portion of its static characteristic. We shall neglect the influence of the coupling of stages through the feed source, and consider that the feed source is sufficiently powerful.

Dynamic Characteristics of Single-Stage Magnetic Amplifiers

Magnetic amplifiers with internal feedback (self-saturation) are characterized by the separation in time of the process of core induction variation caused by the signal action – the control half-period – from the core "interrogation" process – the operating half-period. During the operating half-period, a voltage corresponding to the core magnetization level in the preceding control half-period is produced at the amplifier output. In other words, the output voltage lags behind the input signal by one half-period of the feed voltage.

The output voltage can be approximately characterized by averaging its actual values for each feed-voltage half-period T/2:

$$U_{0n} = \frac{2}{T} \int_{nT/2}^{(n+1)T/2} u_0(t)\,dt \quad (n = 0, 1, 2, 3, \ldots). \tag{1}$$

In this, the amplifier transfer characteristic will be given by a step curve (Fig. 1, Curve 2). The envelope (Curve 1) of the step curve represents the transfer characteristic of the amplifier control circuit.

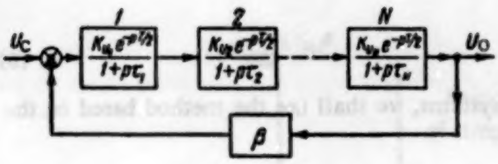


Fig. 2.

The amplifier transfer function which would take into account the discrete character of the output voltage variation can be obtained by writing the amplifier difference equation, and by applying the discrete Laplace transformation to this equation [1,2]. However, in considering magnetic amplifiers with deep negative feedback, we can use only an approximate representation of the transfer function by utilizing the fact that, after feedback is provided, the amplifiers become basically fast-acting. It is well known that fast amplification stages have a lag equal to the feed-voltage half-period. Therefore, the dynamic properties of a stage can be given by the approximate transfer function

$$W_j(p) = \frac{\overline{U}_O(p)}{\overline{U}_C(p)} = \frac{K_{Uj}e^{-pT/2}}{1 + p\tau_j}, \tag{2}$$

where τ_j is the control circuit time constant, which, under steady-state conditions, is related to the voltage gain K_{Uj} of this circuit by the well-known relation* [1]

$$\tau_j = K_{Uj}\frac{w_{Cj}}{2\eta_j f w_{\sim j}}. \tag{3}$$

Here, η_j is the amplifier load circuit efficiency, and w_{Cj} and $w_{\sim j}$ are the numbers of turns in the control and the operating coils, respectively; $K_{Uj} = U_O/U_C$ is determined by taking into account the voltage drop across the signal source internal resistance; f is the feed frequency.

If $\tau_j \gg T$ for individual stages, the step transfer characteristic of a stage can be approximated by the transfer characteristic of the amplifier control circuit, which is shifted by $T/4$ (Fig. 1, Curve 3). Then the approximate transfer function will be given by

$$W_j(p) = \frac{K_{Uj}e^{-pT/4}}{1 + p\tau_j}.$$

The maximum possible lag in the load circuit was used in Eq. (2). Therefore, the approximation used can result in a certain increase in the stability margin.

Method of Investigating the Stability of Multistage Magnetic Amplifiers

The transfer function of an N-stage magnetic amplifier without feedback is determined by the expression

$$W(p) = K_U e^{-pT_l}\prod_{j=1}^{N}\frac{1}{1 + p\tau_j}, \tag{4}$$

where

$$T_l = \sum_{j=1}^{N} T_j \quad \text{and} \quad K_U = \prod_{j=1}^{N} K_{Uj}.$$

Here, T_j and τ_j are the lag and the time constant, respectively, and K_{Uj} is the voltage gain of the jth stage.

If the amplifier is encompassed by rigid negative feedback (Fig. 2), with the feedback factor β (for instance, with respect to voltage), the amplifier transfer function assumes the following form:

$$K(p) = \frac{W(p)}{1 + \beta W(p)} = \frac{K_U}{\beta K_U + e^{pT_l}\prod_{j=1}^{N}(1 + p\tau_j)}. \tag{5}$$

The necessary and sufficient condition for the stability of a system with feedback is that the characteristic equation of the system must not have roots with positive real parts:

---

*Equation (3) can be rigorously proved for nonreversive magnetic amplifiers; it is valid for the majority of reversive circuits. However, if, in the case of any complex circuit, there are doubts whether it can be applied to a reversive amplifier as a whole, it can be applied to each single-cycle element separately.

$$1 + \beta W(p) = 0. \tag{6}$$

For determining the stability conditions for closed-loop lag systems, we shall use the method based on the application of frequency criteria [3,4].

Assuming that p = jω in (6), we shall obtain the equation for determining the critical frequency and lag time values corresponding to the stability boundary:

$$\beta W(j\omega) = -1,$$

or, according to (4),

$$\frac{\beta K_U}{\prod_{j=1}^{N} \sqrt{1+\omega^2\tau_j^2}} e^{-j(\omega T_l + \Sigma \operatorname{arc\,tg} \omega\tau_j)} = 1, \tag{7}$$

whence we find

$$\frac{\beta K_U}{\prod_{j=1}^{N} \sqrt{1+\omega^2\tau_j^2}} = 1 \tag{8}$$

and

$$\omega T_l + \sum \operatorname{arc\,tg} \omega\tau_j = (2m+1)\pi \qquad (m = 0, \pm 1, \pm 2, \ldots), \tag{9}$$

or

$$T_l = \frac{\pi(2m+1)}{\omega} - \frac{\Sigma \operatorname{arc\,tg} \omega\tau_j}{\omega} \qquad (m = 0, \pm 1, \pm 2, \ldots), \tag{10}$$

$$\beta K_U = \frac{1}{\prod_{j=1}^{N} \sqrt{1+\omega^2\tau_j^2}}. \tag{11}$$

After performing calculations by using (10) and (11) for different values of ω, we can plot the stability boundary in the form given by $\beta K_U = f(T_l)$.

If the T_l and the K_U values are assigned for a certain given amplifier, the critical value β_{cr} of the feedback factor corresponding to the stability boundary can be determined from the graph obtained in this manner. In this case, the amplifier stability condition will be expressed by the following inequality:

$$\beta < \beta_{cr}. \tag{12}$$

If the T_l and β values are known, the critical gain $K_{U_{cr}}$ of an open-loop amplifier can be determined in a similar way, and the stability condition can be expressed in the following form:

$$K_U < K_{U\mathrm{cr}}. \tag{13}$$

Finally, if the β and K_U values are given, the adduced equations determine the critical lag $T_{l_{cr}}$, and the amplifier will be stable for

$$T_l < T_{l_{cr}}. \tag{14}$$

Any of the above stability conditions can be used equally in calculations. As will be shown below, they all make it possible to relate the stability of magnetic amplifiers to their structural parameters.

Expressions (10)-(14) are valid for amplifiers having any number of stages, as well as for mixed amplifiers, where not all the stages are magnetic. In the latter case, in determining the value of T_l in Eq. (11) and (14), it is obviously necessary to take into account the fact that, for a nonmagnetic stage, the lag can be different from T/2. If it is necessary to take into account the effect of lag T_c in the magnetic-amplifier control circuit (see [1], Section 6B), it is sufficient to add this value to the over-all amplifier lag.

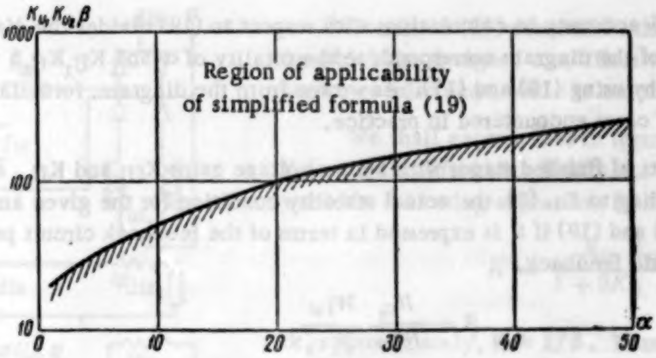


Fig. 3.

As an example, the stability of two-stage magnetic amplifiers will be considered in more detail below. This example is of considerable practical interest because such amplifiers are widely used, and also because, without taking into account the lag, it is impossible to explain the case of unstable operation of these amplifiers, which is observed in practice.

Investigation of the Operating Stability of Two-Stage Magnetic Amplifiers with Negative Feedback

We shall first find the critical feedback factor from Eqs. (10) and (11), which, in the case of a two-stage amplifier, are written as follows:

$$\sqrt{(1+\omega^2\tau_1^2)(1+\omega^2\tau_2^2)} = \beta K_{U_1}K_{U_2} \tag{15}$$

and

$$\omega T_l = \pi(2m+1) - \operatorname{arc\,tg}\omega\tau_1 - \operatorname{arc\,tg}\omega\tau_2 \quad (m = 0, \pm 1, \pm 2, \ldots). \tag{16}$$

The latter equation can be readily reduced to the following form:

$$\omega T_l = \operatorname{arc\,tg}\frac{\omega(\tau_1+\tau_2)}{\omega^2\tau_1\tau_2 - 1}. \tag{17}$$

By solving Eq. (15) with respect to ω^2, and considering that the amplification factors are large, we shall obtain the approximate value

$$\omega^2 = \omega_{cr}^2 \approx \frac{K_{U_1}K_{U_2}\beta}{\tau_1\tau_2}, \tag{18}$$

which, after substitution in (17), makes it possible to obtain the following simplified expression for the critical value of the feedback factor:

$$\beta_{cr} \approx \frac{\tau_1+\tau_2}{K_{U_1}K_{U_2}T_l}, \tag{19}$$

or, if we assume the two-stage amplifier lag time to be $T_l = 1/f$,

$$\beta_{cr} \approx \frac{f(\tau_1+\tau_2)}{K_{U_1}K_{U_2}} \tag{19'}$$

The simplified equation (19) provides a convenient relationship between the amplifier lag, the time constants of individual stages, their gains, and the negative feedback value.

In order to determine the limiting values of the ratio $\alpha = \tau_2/\tau_1$ of the time constants of the stages and the limiting values of the $K_{U_1}K_{U_2}\beta$ product, for which the application of the simplified equation (19) secures a theoretical accuracy of 10% in comparison with calculations where (15) and (17) are used, we performed calculations for the range of α values from 1 to 50, and for the range of $K_{U_1}K_{U_2}\beta$ values from 1 to 2000. The boundary

curve corresponding to a 10% accuracy in calculations with respect to (19) divides the diagram in Fig. 3 into two portions: the lower portion of the diagram corresponds to the totality of α and $K_{U_1}K_{U_2}\beta$ values for which calculations should be performed by using (15) and (17). As we see from the diagram, formula (19) secures sufficient accuracy for the majority of cases encountered in practice.

If the amplifier consists of finished stages with known voltage gains K_{U_1} and K_{U_2}, and the time constants τ_1 and τ_2 determined according to Eq. (3), the actual stability condition for the given amplifier can be determined according to (12) and (19) if β is expressed in terms of the feedback circuit parameters. For instance, for an amplifier with magnetic feedback,

$$\beta = \frac{R_{c_1}}{R_{fw}} \frac{W_{fw}}{W_{c_1}}, \tag{20}$$

where R_{c1} is the resistance of the first-stage control circuit, R_{fw} is the feedback circuit resistance, and W_{fw} and W_{c1} are the number of turns in the feedback windings and in the input stage control winding, respectively. Therefore, we obtain the following stability condition from (12) and (19) for this case:

$$R_{fw} > \frac{K_{U_1} K_{U_2}}{(\tau_1 + \tau_2) f} R_{c_1} \frac{W_{fw}}{W_{c_1}}. \tag{21}$$

By taking (3) into account, this condition can be written in a different form:

$$R_{fw} > \frac{2K_{U_1} K_{U_2} R_{c_1} w_{fw}}{\dfrac{K_{U_1}}{\eta_1 w_{\sim 1}} + \dfrac{K_{U_2} w_{c_2}}{\eta_2 w_{\sim 2} w_{c_1}}}. \tag{22}$$

Expression (22) is convenient for determining the magnetic amplifier stability by using the static characteristics of individual amplifier stages.

Application of Stability Criteria in the Design of Decision Amplifiers

Assume it is necessary to design a decision amplifier with the voltage amplification factor K, where the relative variation of this factor $\gamma = dK/K$, i.e., the amplifier error, does not exceed γ_0. The feed source frequency and voltage, the signal-source resistance, and the load resistance are given.

We shall assume that, from the experience gained in utilization, or as a result of corresponding measurements, the possible variations of the amplification factors of individual stages for certain given operating conditions (a given range of changes in feed voltage, frequency, and temperature) are known, i.e., that $\gamma_1 = dK_{U_1}/K_{U_1}$ and $\gamma_2 = dK_{U_2}/K_{U_2}$ are given. The values of γ_1 and γ_2 are independent of the control circuit parameters.

The initial relationships to be used in designing are obviously given by the inequalities

$$\gamma < \gamma_0, \quad \beta < \beta_{cr}. \tag{23}$$

The amplifier error is related to the errors of the stages by the equation [5]

$$\gamma = \frac{\gamma_1 + \gamma_2}{1 + \beta K_{U_1} K_{U_2}}.$$

As a rule, for a decision amplifier, $\gamma \ll \gamma_1 + \gamma_2$ must hold, which is secured if $\beta K_{U_1} K_{U_2} \ll 1$. In this case,

$$\gamma \approx \frac{\gamma_1 + \gamma_2}{\beta K_{U_1} K_{U_2}}. \tag{24}$$

If we take (24) and (19) into account, conditions (23) assume the following form:

$$\frac{\gamma_1 + \gamma_2}{\beta K_{U_1} K_{U_2}} < \gamma_0, \quad \beta < \frac{\tau_1 + \tau_2}{K_{U_1} K_{U_2} T_l}. \tag{25}$$

By combining these conditions into a single dual inequality, we obtain

$$\frac{\gamma_1 + \gamma_2}{\gamma_0} < \beta K_{U_1} K_{U_2} < \frac{\tau_1 + \tau_2}{T_l}. \tag{26}$$

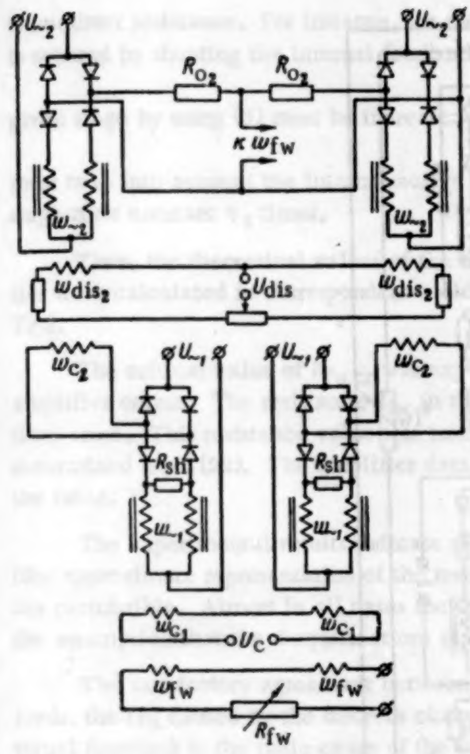


Fig. 4.

Condition (26) determines the range of $\beta K_{U_1} K_{U_2}$ values for which the designed amplifiers will satisfy the necessary accuracy requirements while remaining stable at the same time.

We shall express (26) in terms of the assigned gain value K for an amplifier with feedback. According to (5), we can write the following expression for the gain:

$$K = \frac{K_{U_1} K_{U_2}}{1 + \beta K_{U_1} K_{U_2}},$$

or, approximately, $K \approx 1/\beta$. Hence, we obtain for (26):

$$K \frac{\gamma_1 + \gamma_2}{\beta} < K_{U_1} K_{U_2} < K \frac{\tau_1 + \tau_2}{T_l}. \tag{27}$$

Considering that

$$\tau_1 + \tau_2 = \tau_2 \left(1 + \frac{1}{\alpha}\right) = \frac{K_{U_2} w_{C_2}}{2\eta_2 f w_{\sim 2}} \left(1 + \frac{1}{\alpha}\right),$$

we can obtain the following inequality from (27):

$$\frac{K}{K_{U_2}} \frac{\gamma_1 + \gamma_2}{\gamma_0} < K_{U_1} < K \frac{w_{C_2}}{2\eta_2 w_{\sim 2}} \left(1 + \frac{1}{\alpha}\right); \tag{28}$$

by using this inequality we can determine the allowable range of gain values for one stage with respect to certain chosen parameters of the other stage (in this case, the second stage).

It can happen that the necessary requirements cannot be satisfied by the available amplifier types [in this, inequalities (23) become incompatible]. In this case, it is necessary either to make the requirements less severe with respect to accuracy and/or with respect to the decision amplifier gain, or to provide special measures for improving the amplifier operating stability. Among such expedients are increasing the feed-source frequency, increasing the time constants of the stages (for instance, by means of shorted windings), the introduction of stabilizing (flexible) feedback, etc. It should be noted that improvements in the quality of magnetic materials and valves would make it easier to satisfy conditions (26), which would facilitate the design of stable and highly accurate decision magnetic amplifiers capable of securing considerable gain values.

Experimental Investigation of the Stability of Two-Stage Magnetic Amplifiers

For checking the obtained theoretical results, we used two two-stage reversive amplifier circuits with internal feedback in each stage; the circuits are shown in Figs. 4 and 5. In the first circuit (Fig. 4), which operates at a feed frequency of 500 cps, both stages are based on a differential circuit with the median point in the load circuit. The displacement in the first stage was provided by shunting the feedback rectifier by means of resistor R_{sh}, and separate windings, to which a suitable displacement current was fed, were provided in the second stage.

In the second circuit (Fig. 5), which operates at a feed frequency of 50 cps, the first stage is based on a circuit with ballast resistors, and the second stage is based on a differential circuit with a single common load and commutating resistors in the operating circuits [6].

We tested eight magnetic amplifiers based on the above-described circuits; the amplifiers had different input and operating circuit parameters (see table). For feedback in the first stage of each amplifier, a winding with W_{fw} = 100 turns was provided.

The voltage gains were determined with respect to the static characteristics of the amplifiers, and the time constants of the stages were calculated by using Eq. (3). It should be noted that the time constants must be calculated by taking into account the presence of all additional circuits, which cause an increase in the time constant, as well as by taking into account all the resistances in the control circuit, including the signal-data

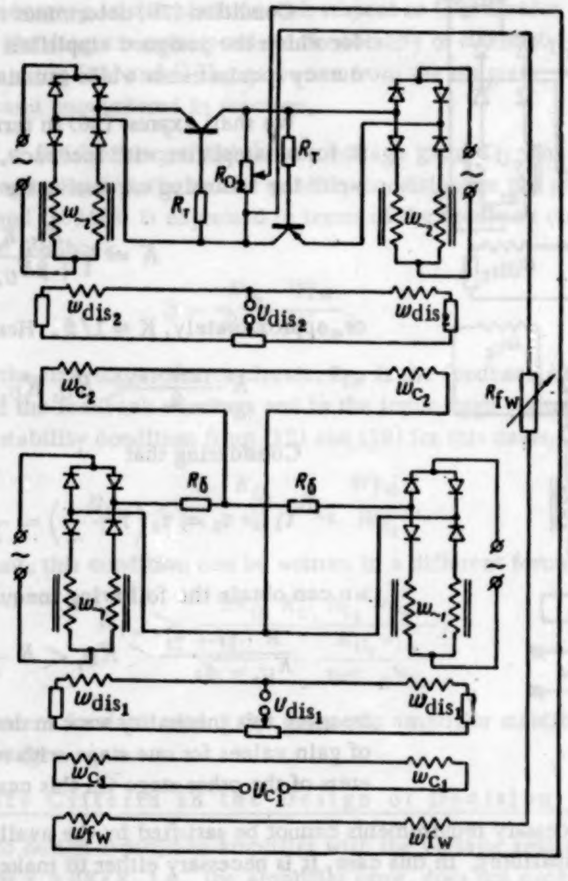


Fig. 5.

Feed frequency, cps	First-stage data							
	Amplifier No.	$w_{\sim_1}$	w_{C_1}	R_{C_1}	R_{O_1}	η_1	K_{U_1}	τ_1
500	1	450	100	19.8	48	0.39	39	0.031
500	2	450	100	65	48	0.39	11.8	0.015
500	3	450	200	39.6	48	0.39	30.8	0.053
500	4	450	200	85.2	48	0.39	11.6	0.041
50	5	450	100	19.8	48	0.39	30	
50	6	450	100	1520	48	0.39	0.024	0.009
50	7	500	1500	223	600	0.17	100	30
50	8	500	1500	500	600	0.17	45	13.5

Feed frequency, cps	Second-stage data						R_{fw}, ohm	
	$w_{\sim_2}$	w_{C_2}	R_{O_2}	η_2	K_{U_2}	τ_2	calculated	measured
500	1500	200	250	0.53	173	0.017	5600	5200
500	1500	200	250	0.53	173	0.017	8500	8600
500	1500	200	250	0.53	208	0.0163	3700	3200
500	1500	200	250	0.53	208	0.0163	4450	4300
50	850	500	250	0.53	625	0.168	3900	3800
50	850	500	250	0.53	625	0.168	4250	4100
50	300	600	50	0.53	83	6	310	270
50	300	600	50	0.53	83	6	575	510

transmitter resistance. For instance, for the circuit shown in Fig. 4, in the first stage of which the displacement is secured by shunting the internal feedback rectifiers by means of resistor R_{sh}, the value of τ_1 obtained for the given stage by using (3) must be increased $\left(1+\frac{w_{\sim 1}^2}{w_{C_1}^2}\frac{R_{C_1}}{R_{sh}}\right)$ times. For the second stages of all amplifiers, one must take into account the internal active resistance of the first-stage output circuit, which reduces the second-stage time constant η_1 times.

Then, the theoretical values of the critical resistance $R_{fw\,cr}$ of the negative feedback circuit in each amplifier were calculated in correspondence with (22). In calculations, it was assumed that the lag of each stage was T/2.

The critical value of $R_{fw\,cr}$ was experimentally determined for the case where no signal was present at the amplifier output. The resistance R_{fw} in the negative feedback circuit was gradually reduced until self-oscillations arose. This resistance value was taken as the critical value, and was compared with the theoretical value determined from (22). The amplifier data and the results obtained by calculation and measurement are given in the table.

The experimental results indicate that the simplifications made in the theoretical analysis of the problem [the approximate representation of the transfer function (1) and the simplifications in deriving (19) and (26)] are permissible. Almost in all cases the theoretical values of $R_{fw\,cr}$ exceed the measured values, which confirms the assumption that the simplifications made lead to a certain improvement in stability.

The satisfactory agreement between the theoretical and the experimental results indicates that, for active loads, the lag caused by the discrete character of changes in the output voltage of magnetic amplifiers with internal feedback is the main cause of the instability of two-stage amplifiers of this type.

SUMMARY

1. In stability investigations, a multistage magnetic amplifier which consists of stages with self-saturation and is provided with negative feedback can be considered as a linear system with lagging feedback. The lag of each stage can be taken with sufficient accuracy to be equal to the feed-voltage half-period.

2. By using the frequency criterion in magnetic amplifier stability analysis, it is possible to obtain a simple relation between the stability condition and the circuit parameters, and to use these conditions in rational design.

3. For active loads, the lag of magnetic amplifiers with self-saturation is the main reason for the excitation of two-stage amplifiers with feedback.

LITERATURE CITED

1. M. A. Rozenblat, Magnetic Amplifiers [in Russian] (Sovet-skoe Radio Press, 1960).
2. P. R. Johannessen, "Analysis of magnetic amplifiers by the use of differential equations," Trans. AIEE 73, 1 (1954).
3. Ya. Z. Tsypkin, "Stability of systems with lagging feedback," Avtomatika i Telemekhanika 7, No. 3 (1946).
4. Ya. Z. Tsypkin, "The degree of stability of systems with lagging feedback," Avtomatika i Telemekhanika 8, No. 3 (1947).
5. F. V. Maiorov, Electronic Regulators [in Russian] (Gostekhizdat, 1956).
6. M. A. Rozenblat and G. V. Subbotina, "Improvement of the efficiency of reversive magnetic dc amplifiers by means of transistors," Avtomatika i Telemekhanika 20, No. 9 (1959).

All abbreviations of periodicals in the above bibliography are letter-by-letter transliterations of the abbreviations as given in the original Russian journal. *Some or all of this periodical literature may well be available in English translation.* A complete list of the cover-to-cover English translations appears at the back of this issue.

SELECTING THE PARAMETER CORRELATIONS OF TWO TYPES OF THIRD-ORDER SINGLE-LOOP AUTOMATIC CONTROL SYSTEMS WITH ADDITIONAL PULSES WITH RESPECT TO THE DERIVATIVE

L. G. Sobolev

Translated from Avtomatika i Telemekhanika, Vol. 22, No. 1,
pp. 107-110, January, 1961
Original article submitted April 20, 1960

Recommendations concerning the selection of parameter correlations of two types of third-order single-loop automatic control systems are given in this article.

1. Statement of the Problem

Third-order single-loop automatic control systems (ACS) with additional pulses with respect to the derivative of the controlled variable are often encountered in practice. In the case of unit step excitation, the transient processes in such systems are determined by the values of four independent parameters while, in most cases, it is desirable that the value of one of these parameters – the system over-all gain – be larger than those of the other three. Investigations performed by several authors [1,2] indicate that, for third-order ACS and unit step excitation, the optimum quality of the transient process is secured if the ACS characteristic equation has triple real roots.

Fig. 1.

2. Graphical Method for Determining Aperiodic Stability Conditions

The characteristic equation of a single-loop ACS consisting of three aperiodic elements with an additional pulse with respect to the derivative has the following form:

$$(p+1)(T_1p+1)(T_2p+1)+\beta p+K=0, \tag{1}$$

where $p \equiv d/d\tau$; T_1 and T_2 are the relative time constants of elements in the control circuit (reduced to the time constant of the system to be controlled), and τ is the relative time (reduced to the time constant of the system to be controlled).

The characteristic equation (1) can be represented by the following system of equations:

$$\begin{aligned} D_1(p) &= (p+1)(T_1p+1)(T_2p+1), \\ D_2(p) &= -\beta p - K, \\ D_1(p) &= D_2(p). \end{aligned} \tag{2}$$

The graphic solution of the system (2) is given in Fig. 1, which, for a certain given value of K, shows the limiting straight lines OA and OB determining the boundaries of aperiodic stability, and the straight line CD determining the maximum value of K_{ap} for aperiodic stability.

It can be readily seen that the values of K_{ap} and β shown in Fig. 1 correspond to the boundary of aperiodic stability pertaining to the case where the characteristic equation has triple real roots.

Thus, the use of this graphic method makes it possible to find the values of β and K_{ap} with respect to certain given T_1 and T_2 values, and also to find the regions of K and β values corresponding to aperiodic stability.

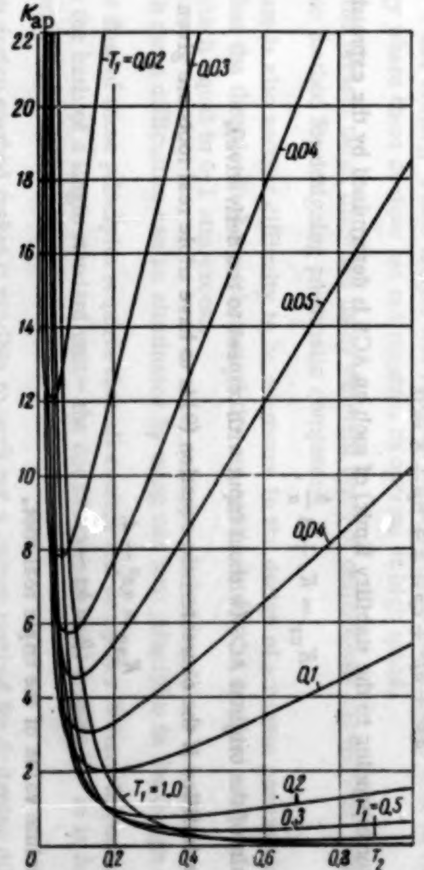


Fig. 2.

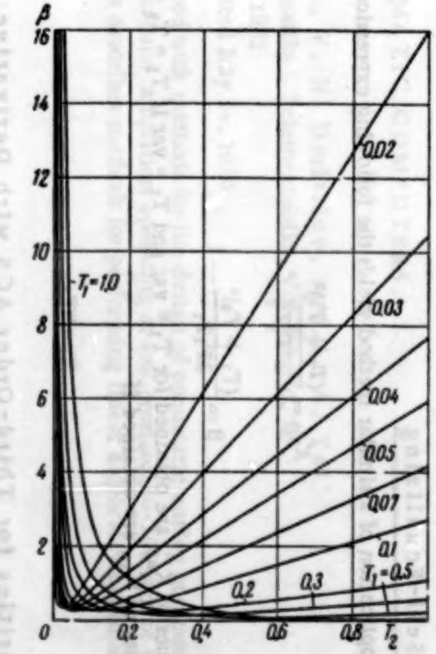


Fig. 3.

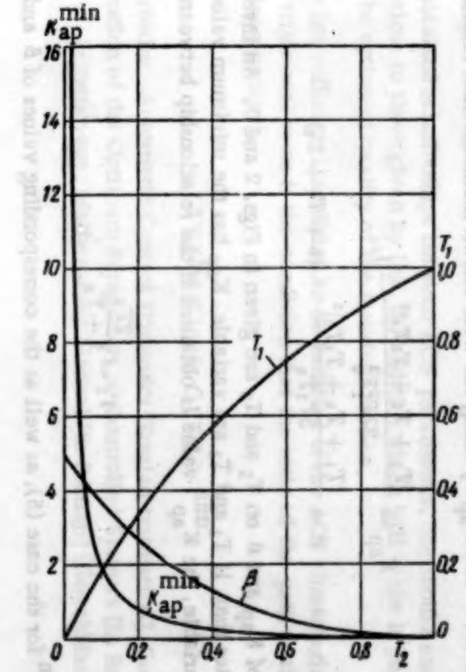


Fig. 4.

3. Conditions for the Extremum of K_{ap}

The straight line CD is tangent to the $D_1(p)$ curve at the inflection point, which makes it possible to find the analytical expressions for β and K_{ap} by using the well-known mathematical relations for third-power polynomials [3]:

$$K_{ap} = \frac{(T_1 + T_2 + T_1T_2)^3}{27T_1^2T_2^2} - 1, \qquad (3)$$

$$\beta = \frac{(T_1 + T_2 + T_1T_2)^2}{3T_1T_2} - (1 + T_1 + T_2). \qquad (4)$$

The dependences of K_{ap} and β on T_1 and T_2 are given in Figs. 2 and 3. An investigation of expression (3) for the extremum indicates that, if T_1 and T_2 are variable, K_{ap} has the minimum value K_{ap}^{min} for $T_1 = T_2 = 1$. If T_2 = const and T_1 is variable, the K_{ap}^{min} value is obtained if the relationship between T_1 and T_2 has the following form:

$$T_1 = \frac{2T_2}{1 + T_2}. \qquad (5)$$

The values of K_{ap}^{min} for the case (5), as well as the corresponding values of β and T_1 in dependence on T_2, are shown in Fig. 4.

4. Selection of Parameter Relations in the Case Where the System to Be Controlled Is Not Self-Equalizing

In this case, the application of a similar method yields the following expressions:

$$K_{ap} = \frac{(T_1 + T_2)^3}{27T_1^2T_2^2}, \qquad (6)$$

$$\beta = \frac{(T_1 + T_2)^2}{3T_1T_2} - 1. \qquad (7)$$

The minimum values of K_{ap} are obtained for T_1 = var and T_2 = var if $T_1 = T_2 = 1$; for T_2 = const and T_1 = = var, the minimum values of K_{ap} are obtained if

$$T_1 = 2T_2. \qquad (8)$$

5. General Regularities for Third-Order ACS with Derivative Input

The above method can be applied to third-order ACS with input with respect to the derivative if the ACS characteristic equation is given by

$$ap^3 + bp^2 + cp + d + \beta p + K = 0. \qquad (9)$$

The critical gain (corresponding to the stability limit) of such an ACS is determined by the expression

$$K_{cr} = K_0 + \beta\frac{b}{a}, \qquad (10)$$

where K_0 is the critical gain of the original ACS (without input with respect to the derivative).

The conditions to be fulfilled if the characteristic equation (9) is to have triple real roots are given by

$$K_{ap} = a\alpha^3 - d, \qquad (11)$$

$$\beta = b\alpha - c, \qquad (12)$$

where $\alpha = b/3a$ is the absolute value of the triple real root.

For the value of β obtained from (12), the following relation holds:

$$\frac{K_{cr} + d}{K_{ap} + d} = 9. \qquad (13)$$

SUMMARY

1. In order to obtain triple real roots of the characteristic equations of the considered ACS types, it is necessary to determine the K_{ap} and β values by using Eqs. (3), (4), (6), and (7).

2. From the various obtained relationships between ACS parameters, one should exclude the relationships between T_1 and T_2 that are close to those given by Eqs. (5) and (8), which will make it possible to improve the static accuracy of ACS for the optimum quality of the transient process.

3. The necessary (but not sufficient) condition to be fulfilled if the ACS characteristic equation of the type given by (9) is to have triple real roots is the condition that the ratio of the gain to the critical gain must be equal to 9 (for a closed-loop system).

LITERATURE CITED

1. R. Oldenburg and G. Sartorius, Automatic Control Dynamics [Russian translation] (Gosénergoizdat, 1949).
2. V. A. Bodner, Determination of the Optimum Parameters of Controlled Systems [in Russian] (Oborongiz, 1953).
3. I. N. Bronshtein and K. A. Semendyaev, Mathematics Manual [in Russian] (Gostekhizdat, 1955).

A NEW TYPE OF PNEUMATIC COMPUTERS. I

G. T. Berezovets, V. N. Dmitriev, and A. A. Tal'

Translated from Avtomatika i Telemekhanika, Vol. 22, No. 1, pp. 111-118, January, 1961

Original article submitted July 24, 1960

This article is concerned with methods for the design of pneumatic analog computers that are based on pneumatic decision amplifiers operating at low pressures (0-100 mm of water). The first part of this article describes methods for performing linear algebraic operations.

1. Introduction

Earlier, the development of industrial pneumatic automation followed only along the lines of designing various regulating devices and servomechanisms. In recent years, communications on the development of pneumatic analog and digital computers have been appearing more and more frequently in our country and abroad [1-6 and others]. These simple and reliable devices can be successfully used in the rather large fields of technology where these devices are not required to operate at high speed.

At the present time, there are many different types of components and devices, which is due to the absence of a unique method for designing pneumatic computers.

There is also another difficulty to be overcome in the design of accurate pneumatic computers, namely, the fact that the throttle discharge characteristics are nonlinear for the standard range of input and output pressures, which is equal to 0-1 atm excess.

Both these difficulties can be eliminated by using two new principles in the design of pneumatic devices.

The first of these principles requires that all control and computer devices in pneumatic automation be designed on the basis of a single standard unit – the decision (operating) amplifier – as is done in the case of electronic computers.

This design method makes it possible to work out a unique method for designing different types of pneumatic devices, and to minimize the number of necessary parts.

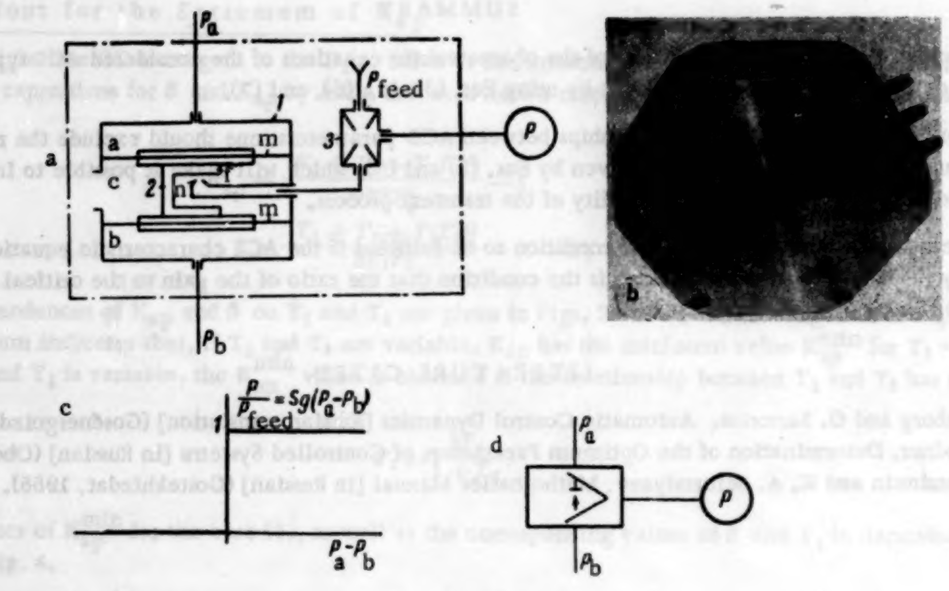


Fig. 1.

The second of these two new principles requires that the range of operating pressures be only one-hundredth of the range presently used, i.e., that pressures in the 0-100 mm of water range be used.*

The switch to low operating pressures does not entail a deterioration of the dynamic characteristics of these devices; it provides a solution to the problem of designing linear throttles, and the air consumption is thereby reduced at least tenfold.

The main purpose of the present article is to demonstrate how most diverse computers can be designed by using pneumatic decision amplifiers operating in the 0-100 mm of water pressure range. The present article is concerned with different components for this system and methods for performing linear algebraic operations. The methods for the design of integrating and differentiating computers, as well as computers for some nonlinear operations, will be given in the second part of the article.

Striving to encompass the greatest number of different computer operations, the authors will here consider only layout problems related to the operating principles of computers without dealing either with technical problems connected with the construction of the devices, or problems involving their operating accuracy.

2. The System Components

According to the proposed principles for the design of pneumatic computers of the type under consideration, standard parts are used in the layout of each computer. The various parts used in the system can be divided into two groups. The first group pertains to the basic unit in the system, the pneumatic decision amplifier, and the second group comprises pneumatic throttles and pneumatic capacitors ordinarily used in pneumatic automatic devices. Moreover, individual units used in pneumatic relay technology (which are also standardized) (see [5]), can be applied in devices of this type.

The pneumatic decision amplifier, whose layout and external view are shown in Fig. 1, consists of a frame 1 with a fixed nozzle n; a membrane unit 2, with two identical thin membranes m, the rigid centers of which are connected by means of a bracket f, which acts as a flapper in front of the nozzle; and a permanent throttle 3 which can be pulled out. The membrane unit is placed inside frame 1, whereby three chambers a, b, and c are formed. Pressures P_a and P_b are fed to chambers a and b, respectively, and chamber c, where the nozzle and the flapper are mounted, is open to atmosphere. One side of the pull-out throttle is connected to the feed line, where

* The idea of designing devices operating at low pressures belongs to Ferner (East Germany). On the basis of this principle, he constructed devices which yielded satisfactory results [1].

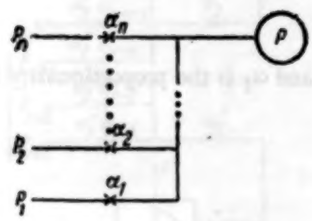


Fig. 2.

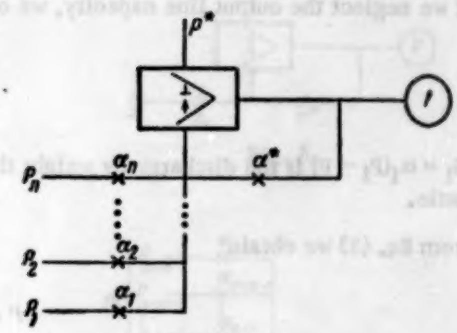


Fig. 3.

the pressure is equal to P_{feed}, and the other side is connected to nozzle n. For the amplifier, P_a and P_b are input pressures, and the pressure P, which is established in the pull-out throttle chamber, provides the amplifier output signal. It is obvious that $P = P_{feed}$ when the flapper completely covers the nozzle, and that P = 0 atm ex if the nozzle is not completely covered by the flapper. This is due to the fact that, instead of an ordinary, a pull-out nonvariable, throttle is used in this amplifier.

It can be readily seen that the following function can be simulated in this decision amplifier in correspondence with its layout:

$$\frac{P}{P_{feed}} = \mathrm{Sg}\,(P_a - P_b), \tag{1}$$

i.e., $P = P_{feed}$ if $P_a > P_b$ and $P = 0$ if $P_a < P_b$ (Sg is a function equal to 1 if $P_a - P_b > 0$; it is equal to 0 if $P_a - P_b < 0$). The static characteristic shown in Fig. 1c illustrates this dependence. The greater the accuracy with which function (1) is simulated in the decision amplifiers, the greater the accuracy of devices based on these amplifiers.

In other layouts, the decision amplifier will be conventionally represented in the form shown in Fig. 1d.

Three types of pneumatic throttles are used in this system:

1) a nonvariable linear (laminar) throttle, whose static discharge factor is given by $G = \alpha\Delta p$, where G is the discharge by weight through the throttle, Δp is the pressure drop in the throttle, and α is a constant proportionality factor;

2) a nonvariable quadratic (turbulent) throttle, whose discharge factor is given by $G = \beta\sqrt{\Delta p}$;

3) a variable linear throttle, whose discharge factor is the same as that of the corresponding nonvariable throttle, and which differs from the latter only by the variability of the α coefficient.

3. Simulation of Linear Algebraic Functions

In order to perform any linear algebraic operation by means of pneumatic devices, one must have general means for simulating a dependence of the following form:

$$P = \sum_{i=1}^{n} k_i P_i \qquad (-\infty < k_i < +\infty), \tag{2}$$

where P_i are the input pneumatic signals (excess pressure), and P is the output pressure.

Let us consider, in turn, three methods by means of which this problem can be solved in various degrees.*

The First Method. Figure 2 shows the schematic layout of a noncompensating throttle summation device. In this layout, the input pressures $P_1, P_2, \ldots, P_n$ are fed to nonvariable linear throttles $\alpha_1, \alpha_2, \ldots, \alpha_n$. The output pressure P is established in the line connecting the outputs of all throttles.

*Each of these methods corresponds to an electronic method which has been applied to pneumatics.

If we neglect the output line capacity, we obtain

$$\sum_{i=1}^{n} G_i = 0, \tag{3}$$

where $G_i = \alpha_i(P_i - P)$ is the discharge by weight through the _ith throttle, and α_i is the proportionality factor of this throttle.

From Eq. (3) we obtain

$$P = \sum_{i=1}^{n} k_i P_i, \tag{4}$$

where

$$k_i = \frac{\alpha_i}{\sum_{i=1}^{n} \alpha_i},$$

and, consequently,

$$0 < k_i < 1, \quad \sum_{i=1}^{n} k_i = 1. \tag{5}$$

Conditions (5) considerably restrict the capability of the summation device under consideration: it cannot be used for ordinary addition, since $k_i < 1$; neither can it be used for subtraction, since $k_i > 0$.

However, summation devices of this type, being inordinately simple, are widely used in pneumatic automation devices for the solution of problems satisfying conditions (5). As an example, we can indicate the summation device of the positive feedback chamber in the ACS control unit [7].

The Second Method. Figure 3 shows the layout of a compensating inverting throttle summation device, which is based on the application of a decision amplifier.

The layout consists of n + 1 linear throttles ($\alpha_1, \alpha_2, \ldots, \alpha_n$ and α^*) with connected outputs and a single decision amplifier. Input pressures $P_1, P_2, \ldots, P_n$ are fed to n throttles ($\alpha_1, \alpha_2, \ldots, \alpha_n$), and pressure P, which arises at the decision amplifier output, is fed to the input of the (n + 1)th throttle (α^*). The constant pressure $P_a = P^* = 50$ mm of water is maintained in the chamber a of the decision amplifier, and the chamber b is connected by means of the line connecting the outputs of all the n + 1 throttles. If we take into account the fact that the presence of the negative feedback shunting the decision amplifier secures the maintenance of the pressure $P_b = P^*$ in the b chamber, and if we neglect the capacities of the connecting line and of the decision amplifier b chamber, we obtain

$$\alpha^* (P - P^*) + \sum_{i=1}^{n} \alpha_i (P_i - P^*) = 0. \tag{6}$$

By introducing the notation

$$\overline{P}_i = P_i - P^*, \quad \overline{P} = P - P^*, \tag{7}$$

i.e., by measuring all pressures from the "conventional zero P^*," we obtain

$$\overline{P} = -\sum_{i=1}^{n} k_i \overline{P}_i, \tag{8}$$

where

$$k_i = \frac{\alpha_i}{\alpha^*}.$$

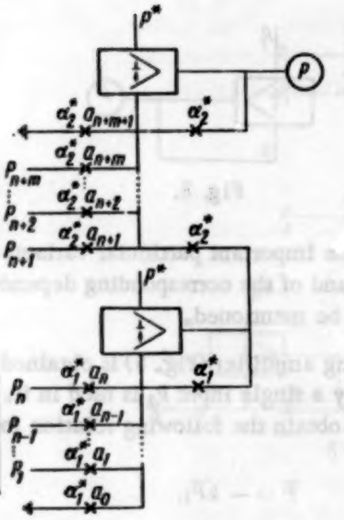


Fig. 4. Layout of the device capable of performing any linear algebraic operation by means of two decision amplifiers.

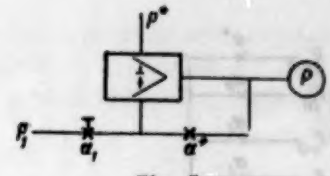


Fig. 5.

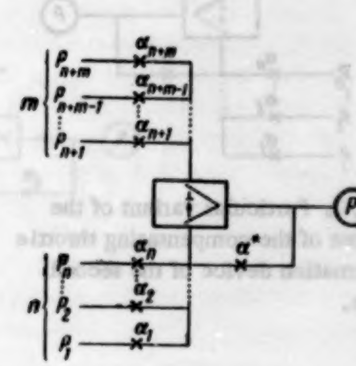


Fig. 6.

In this case, the value of k_i varies within the interval

$$0 < k_i < \infty. \tag{9}$$

Any linear algebraic function can be simulated by using these summation devices. In order to demonstrate this, let us consider the dependence (2) when written in the following form:

$$P = \sum_{i=1}^{n} a_i P_i - \sum_{i=1}^{m} a_{n+i} P_{n+i}. \tag{10}$$

If notation (7) is used, the dependence (10) assumes the following form:

$$\overline{P} = \sum_{i=1}^{n} a_i \overline{P}_i - \sum_{i=1}^{m} a_{n+i} \overline{P}_{n+i} + \left[\sum_{i=1}^{n} a_i - \left(1 + \sum_{i=1}^{m} a_{n+i} \right) \right] P^*. \tag{11}$$

Moreover, considering that, if any $P_i = 0$, $\overline{P}_i = -\overline{P}^*$, the dependence (11) will be written thus:

$$\overline{P} = \sum_{i=0}^{n} a_i \overline{P}_i - \sum_{i=1}^{m+1} a_{n+i} \overline{P}_{n+i}, \tag{12}$$

where

$$a_0 = 1 + \sum_{i=1}^{m} a_{n+i}, \quad a_{n+m+1} = \sum_{i=1}^{n} a_i, \quad P_0 = P_{n+m+1} = 0. \tag{13}$$

Relation (12) can be written in its final form:

$$\overline{P}_a = -\sum_{i=0}^{n} a_i \overline{P}_i, \qquad \overline{P} = -\left(\overline{P}_a + \sum_{i=1}^{m+1} a_{n+i} \overline{P}_{n+i} \right). \tag{14}$$

The layout shown in Fig. 4 can simulate the dependence (14) and, consequently, the initial function (10). This confirms the possibility of obtaining any linear algebraic function by means of this layout if, in the general case, two decision amplifiers are used.

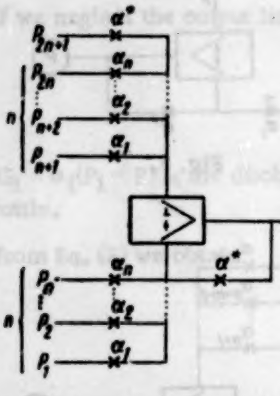


Fig. 7. Particular variant of the layout of the compensating throttle summation device of the second type.

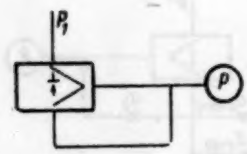


Fig. 8.

From among the important particular variants of the layout given in Fig. 3, and of the corresponding dependence (8), the following should be mentioned.

A. An inverting amplifier (Fig. 5) is obtained from the general layout if only a single input P_1 is used in it. Assuming that n = 1 in (8), we obtain the following relation for this layout:

$$\overline{P} = -k\overline{P}_1, \tag{15}$$

where

$$k = \frac{\alpha_1}{\alpha^*}.$$

B. We obtain an inverter instead of the inverting amplifier if

$$k = \frac{\alpha_1}{\alpha^*} = 1.$$

In this case, we have

$$\overline{P} = -\overline{P}_1. \tag{16}$$

The Third Method. Figure 6 shows the layout of a compensating throttle summation device of the second type. It differs from the previous type by the fact that the chamber a of the decision amplifier is here connected to the output of a noncompensating throttle summation device with m inputs. By using relation (4) for determining the pressure P_a that arises in the decision amplifier chamber a, we obtain

$$P_a = \frac{1}{\sum_{i=1}^{m} \alpha_{n+i}} \sum_{i=1}^{m} \alpha_{n+i} P_{n+i}. \tag{17}$$

In the same manner, for the amplifier b chamber, we have

$$P_b = \frac{1}{\alpha^* + \sum_{i=1}^{n} \alpha_i} \left(\alpha^* P + \sum_{i=1}^{n} \alpha_i P_i \right). \tag{18}$$

The presence of compensating feedback in the layout (the decision amplifier output acts on pressure P_b through the α^* throttle) leads to the fact that $P_a = P_b$ and, consequently,

$$\frac{1}{\alpha^* + \sum_{i=1}^{n} \alpha_i} \left(\alpha^* P + \sum_{i=1}^{n} \alpha_i P_i \right) = \frac{1}{\sum_{i=1}^{m} \alpha_{n+i}} \sum_{i=1}^{m} \alpha_{n+i} P_{n+i}. \tag{19}$$

Hence,

$$P = q \sum_{i=1}^{m} \alpha_{n+i} P_{n+i} - \sum_{i=1}^{n} \alpha_i P_i. \tag{20}$$

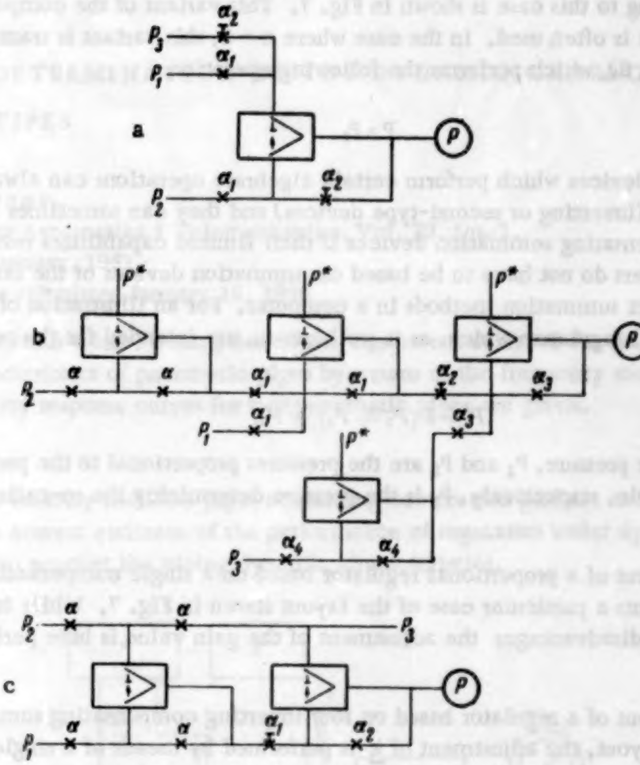


Fig. 9

where

$$a_i = \frac{\alpha_i}{\alpha^*}, \quad q = \frac{1 + \sum_{i=1}^{n} a_i}{\sum_{i=1}^{m} a_{n+i}}.$$

It follows from relation (20) that the summation device under consideration makes it possible to perform any linear algebraic operation given by (10) by using only a single decision amplifier, in contrast to the summation device based on the previous layout, where at least two decision amplifiers had to be used for this purpose.

In particular, it also follows from (20) that, if

$$m = n + 1, \quad \alpha_i = \alpha_{n+i} \quad (i = 1, \ldots, n) \quad \text{and} \quad \alpha_{2n+1} = \alpha^*, \tag{21}$$

then

$$q = 1, \quad a_i = a_{n+i} \quad (i = 1, \ldots, n) \quad \text{and} \quad a_{2n+1} = 1, \tag{22}$$

and, consequently,

$$P = P_{2n+1} + \sum_{i=1}^{n} a_i (P_i - P_{n+i}), \tag{23}$$

where

$$a_i = \frac{\alpha_i}{\alpha^*}.$$

The layout corresponding to this case is shown in Fig. 7. This variant of the compensating throttle summation device of the second type is often used. In the case where n = 0, this variant is transformed into the layout of the so-called repeater (Fig. 8), which performs the following operation:

$$P = P_1. \tag{24}$$

Pneumatic automation devices which perform certain algebraic operations can always be based on compensating summation devices (inverting or second-type devices) and they can sometimes be constructed on the basis of very simple noncompensating summation devices if their limited capabilities resulting from this are acceptable. In this, the computers do not have to be based on summation devices of the same type; it is sometimes advisable to use different summation methods in a computer. For an illustration of this, we shall consider possible layouts of proportional regulators which, as is well known, are intended for the performance of the following operation:

$$P = k(P_1 - P_2) + P_3, \tag{25}$$

where P is the regulator output pressure, P_1 and P_2 are the pressures proportional to the present and the assigned values of the controlled variable, respectively, P_3 is the pressure determining the so-called control point, and $\underline{k}$ is the regulator gain.

Figure 9a shows the layout of a proportional regulator based on a single compensating summation device of the second type. It represents a particular case of the layout shown in Fig. 7. While it is extremely simple, this layout has a considerable disadvantage: the adjustment of the gain value is here performed by means of two pneumatic throttles.

Figure 9b shows the layout of a regulator based on four inverting compensating summation devices (two of them are inverters). In this layout, the adjustment of $\underline{k}$ is performed by means of a single pneumatic throttle; however, this results in a much more complicated layout. Finally, Fig. 9c shows a regulator layout with a single compensating inverting summation device and a single compensating summation device of the second type. Here, the gain is adjusted by means of a single throttle, and, moreover, this layout is much more efficient that that shown in Fig. 9b from the viewpoint of the amount of equipment used. The above example indicates that, in certain cases, the application of combined summation devices of different types is the most efficient procedure.

LITERATURE CITED

1. W. Ferner, "Pneumatic regulator units," Problems in Pneumatic and Hydraulic Automation [in Russian] (AN SSSR Press, 1960).
2. Yu. I. Ivlichev and É. M. Nadzhafov, "Problems in pneumatic computer design," Problems in Pneumatic and Hydraulic Automation [in Russian] (AN SSSR Press, 1960).
3. Yu. I. Ivlichev, "A universal pneumatic multiplier-divider and a device for square-root evolution," Avtomatika i Telemekhanika 19, No. 11 (1958).
4. N. D. Lanin, "A small-size pneumatic analog computer and the lag unit," Problems in Pneumatic and Hydraulic Automation [in Russian] (AN SSSR Press, 1960).
5. T. K. Berends and A. A. Tal', "Pneumatic relay circuits," Avtomatika i Telemekhanika 20, No. 11 (1959).
6. G. T. Berezovets, A. L. Malyi, and É. M. Nadzhafov, Component Devices of Standard Pneumatic Units and Their Application in the Automation of Industrial Processes [in Russian] (Gostoptekhizdat, 1960).
7. L. A. Zalmanzon and A. I. Semikova, "Investigation of the characteristics of pneumatic chambers which are used as summation devices," Problems in Pneumatic and Hydraulic Automation [in Russian] (AN SSSR Press, 1960).

All abbreviations of periodicals in the above bibliography are letter-by-letter transliterations of the abbreviations as given in the original Russian journal. *Some or all of this periodical literature may well be available in English translation.* A complete list of the cover-to-cover English translations appears at the back of this issue.

EXPERIMENTAL DETERMINATION OF THE DYNAMIC CHARACTERISTICS OF PNEUMATIC PIPES

A. M. Smirnov

Translated from Avtomatika i Telemekhanika, Vol. 22, No. 1, pp. 119-120, January, 1961
Original article submitted January 18, 1960

This article provides a brief description of the equipment designed for the determination of the dynamic characteristics of pneumatic pipes by means of the frequency method. The experimental frequency response curves for two pneumatic pipes are given.

Pneumatic pipes (usually metallic pipes containing compressed gas) are used in pneumatic regulators and control systems. For a correct estimate of the performance of regulators under dynamic operating conditions, it is necessary to take into account the piping dynamic characteristics.

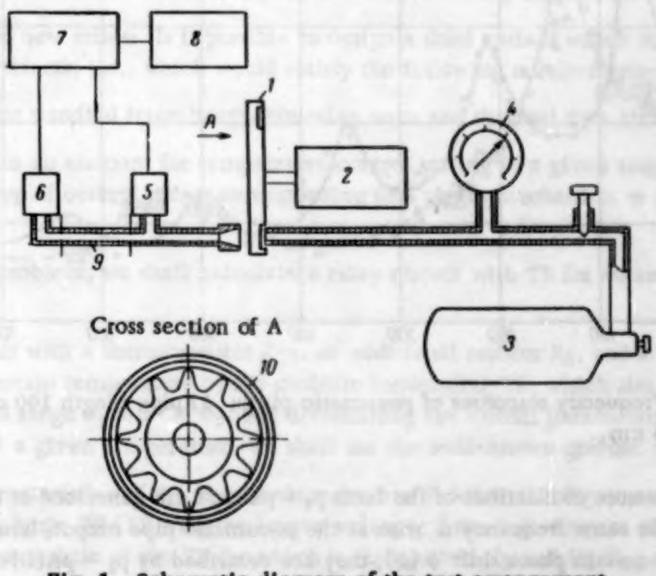


Fig. 1. Schematic diagram of the test arrangement.

An exact theoretical determination of these characteristics is a very complicated problem. Therefore, an experimental method for determining the frequency characteristics of pneumatic pipes is given below.

For the experimental determination of the frequency characteristics of pneumatic pipes, it is necessary to have equipment for generating harmonic pressure oscillations at the pneumatic pipe input and for measuring and recording pressure oscillations at the pneumatic pipe input and output without significant dynamic errors.

The test arrangement schematically shown in Fig. 1 can be used for this purpose. The stand consists of a special device for generating pressure oscillations and an inductive two-channel apparatus for measuring the pressure oscillations.

The device for generating pressure oscillations consists of the following parts: a disk 1 with profiled openings, which is driven by an electric motor 2; a compressed-air tank 3; a control pressure gauge 4. The stream of compressed air supplied from the tank to the input of the pneumatic pipe under investigation, 9, is interrupted by disk 1, due to which near-sinusoidal pressure oscillations arise in the pneumatic pipe. This device makes it possible to generate pressure oscillations with frequencies of up to 1000 cps and an oscillation depth (double amplitude) of up to 1 atm for a pressure of 15 atm in the pneumatic feed pipe.

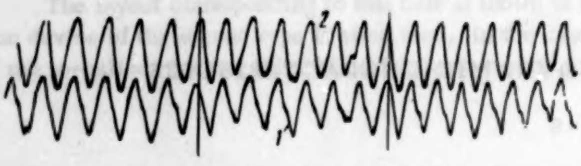


Fig. 2. Sample of the recording of pressure oscillations with a frequency of 70 cps. 1) Pneumatic pipe inlet; 2) pneumatic pipe outlet.

The rotating disk has a diameter of 18 cm and is provided with 12 profiled through-cuts. The inside diameter of the pressure-feed pipe and of the receiving pipes is equal to 6 mm. A specimen of pressure recording is shown in Fig. 2. The two-channel inductive apparatus consists of two inductive pressure-data transmitters 5 and 6, which are mounted at the pneumatic pipe input and output, an amplifier-converter 7, and a loop oscillograph 8. This equipment makes it possible to measure and regulate pressure oscillations with frequencies of up to 1000 cps with a frequency response error not greater than 5%.

The inductive data transmitters have a small pressure chamber, the volume of which does not exceed 0.5 cm^3, due to which the volumes of the pressure chambers of the data transmitters that are connected to the pneumatic pipe can be neglected.

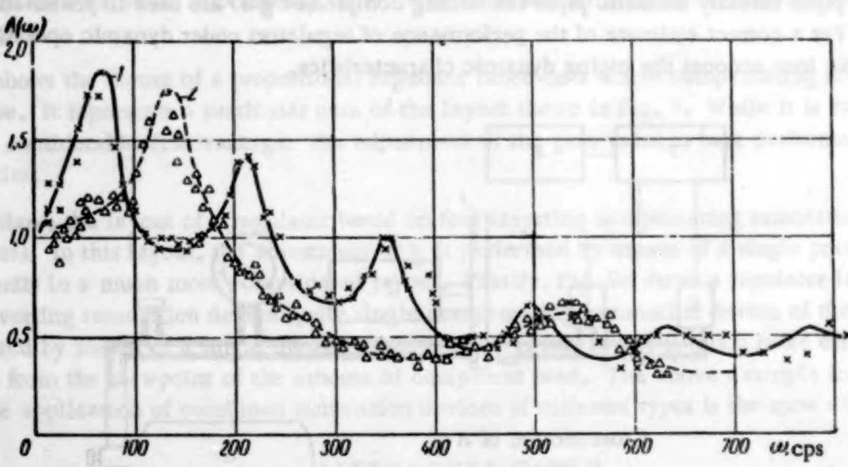


Fig. 3. Frequency responses of pneumatic pipes: 1) pipe length 100 cm; 2) pipe length 25 cm.

Thus, harmonic pressure oscillations of the form $p_1 = p\cos\omega t$ are generated at the pneumatic pipe input. Pressure oscillations of the same frequency ω arise at the pneumatic pipe output; however, they have a different amplitude, $pA(\omega)$, and a certain phase shift $\varphi(\omega)$; they are described by $p_2 = pA(\omega)\cos[\omega t - \varphi(\omega)]$. By comparing the ordinates on the pressure oscillation recordings at the pneumatic pipe input and output, the frequency response

$$A(\omega) = f(\omega)$$

is obtained on the oscillograph strip chart.

Figure 3 shows the amplitude responses of two 4 × 6 mm copper pneumatic pipes with lengths equal to 100 and 25 cm, which were determined experimentally according to the above method.

From the data in Fig. 3, it is clear that the frequency responses of pneumatic pipes in the frequency range of up to 700 cps have a number of resonance maxima: The pneumatic pipe 100 cm long has five maxima, and the pneumatic pipe 25 cm long has two maxima. The largest frequency response value occurs at the first resonance maximum, where it attains 1.7-1.8. For the subsequent resonance maxima, the amplitude response values gradually decrease and, at frequencies close to 600 cps, the amplitude response is equal to 0.5-0.6.

THE DESIGN OF RELAY CIRCUITS WITH THERMORESISTORS

V. F. Bakhmut-skii

Translated from Avtomatika i Telemekhanika, Vol. 22, No. 1,
pp. 121-124, January, 1961
Original article submitted June 6, 1960

A method for designing relay circuits with thermoresistors for an assigned temperature control range is described. Practical recommendations and formulas are given.

1. Temperature control devices where the relay effect is used in circuits with a thermoresistor (TR) are widely used in industry because they are simple and reliable.

At the present time, two basic variants of these devices have been developed:

1) devices with small units and fixed temperature control setting (for instance, UTS-1, [1], p. 166);

2) devices with large units with variable temperature control setting (for instance, UTK-6 [1]).

The following question now arises: Is it possible to design a third variant which would combine the advantages of the above two variants, i.e., which would satisfy the following requirements:

a) The device contains standard interchangeable relay units and thermal data transmitters.

b) The device contains an element for temperature control setting in a given range (and, correspondingly, a setting scale) for each group of control points corresponding to a single mechanism or to a single portion of the technological cycle.

In order to solve this problem, we shall calculate a relay circuit with TR for an assigned temperature control range.

2. Consider the circuit with a thermoresistor R_{T1}, an additional resistor R_a, and a correcting resistor R_c (see figure). Let θ^* be a certain temperature of the medium surrounding TR, which lies within the limits of the assigned temperature control range $\theta_1 \le \theta^* \le \theta_2$. For determining the circuit parameters corresponding to the advent of the relay effect at a given temperature, we shall use the well-known graphic plot [1].

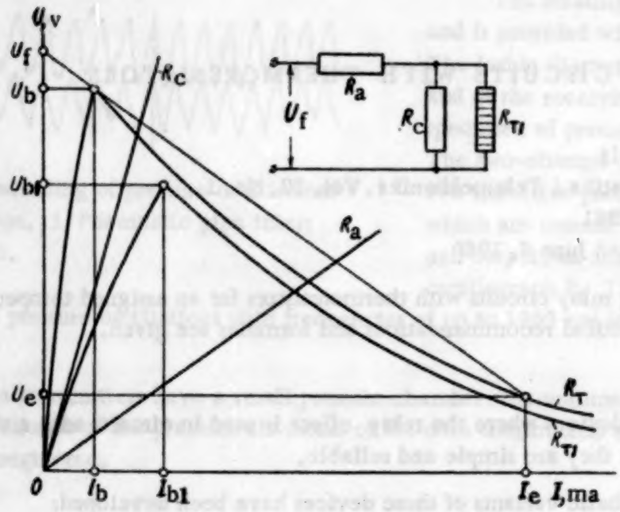
The volt-ampere characteristics of the circuit elements are shown in the figure, where R_T denotes the volt-ampere characteristic of the basic TR (TR with the largest voltage of the volt-ampere characteristic maximum), R_{T1} is the volt-ampere characteristic of the TR for which R_c is selected, I_b and U_b (I_{b1} and U_{b1}) are the current in, and the voltage across, R_T (R_{T1}) at the beginning of the relay effect, I_e and U_e are the current in, and the voltage across, R_T at the end of the relay effect, and U_f is the circuit feed voltage. The value of I_e is determined by the relay actuation current $I_e \ge I_{av}$ under the condition that the relay winding resistance not exceed R_a. The values of R_a, U_f, and R_c can be found graphically or by calculation according to Eqs. (123), (124), and (133) [1]:

$$R_a(\theta^*) = \frac{U_b(\theta^*) - U_e(\theta^*)}{I_e - I_b(\theta^*)}, \tag{1}$$

$$U_f(\theta^*) = U_b(\theta^*) + I_b(\theta^*)\,R_a(\theta^*), \tag{2}$$

$$R_c(\theta^*) = \frac{U_{b1}(\theta^*)\,R_a(\theta^*)}{U_b(\theta^*) - U_{b1}(\theta^*)}. \tag{3}$$

If the indicated calculation method is applied to the entire $\theta_1 \le \theta \le \theta_2$ range, we shall obtain different values of R_a, U_f, and R_c for different θ values; in this case, for switching over from one controlled temperature to another, it will be necessary to change three parameters simultaneously, which is not acceptable. The problem consists in changing one parameter so that the other two remain constant. Since R_c depends on R_a, the values of these two resistances should be kept constant, and the temperature setting within the given range should



be secured by varying U_f; with such an arrangement, the relay units and the thermal data transmitters will be interchangeable, i.e., the requirement a in paragraph 1 will be satisfied; it is obvious that the requirement b will also be fulfilled.

We shall now determine the actual values of R_a and R_c, and the dependence $U_f(\theta)$. Considering that the minimum value of I_e corresponds to the maximum temperature, we shall write

$$R_a(\theta) = R_a(\theta_2) = \frac{U_b(\theta_2) - U_e(\theta_2)}{I_e - I_b(\theta_2)} . \tag{4}$$

The determination of R_a is restricted by the following requirements: TR must not be heated to a temperature exceeding the maximum allowable temperature; the long-duration current passing through the relay winding after it has been actuated must not exceed the maximum allowable value. The first requirement is satisfied by choosing a suitable relay [for TR of the KMT-11 type, $I_{av}U_{av}(\theta_1) \leq 0.2$ w must hold] and by blocking TR when it is actuated. The simplest way of meeting the second requirement is to provide an arrangement for disconnecting the circuit feed after actuation or all the circuits of a given group pertaining to a single controlled mechanism. In the case where it is necessary to change the setting and to continue the controlling operation after the circuit has been actuated, the feed is temporarily disconnected. The time during which the feed is to be disconnected is determined experimentally from the condition that there should be no spurious actuation: This time must be sufficient for TR to cool down (approximately) to the surrounding medium temperature after it has been heated to 300-400°C at the moment of actuation; delay time must be effected automatically by means of a time relay. Besides disconnecting the relay, its overloading can be prevented by other means. For instance, in the case where the number of changes in the feed voltage $U_f(\theta)$ in the $\theta_1 \leq \theta \leq \theta_2$ range is close to (or less than) the number of times the actuation and dropout currents pass through the relay, it is advisable to introduce ballast resistors, which are connected in parallel or in series to the relay winding when it is actuated; this would make it possible to reduce the relay winding current to the maximum allowable value for the largest feed voltage $U_f(\theta_1)$ while keeping its value larger than the relay dropout current for the minimum feed voltage $U_f(\theta_2)$.

The values of $U_f(\theta)$, where $\theta_1 \leq \theta \leq \theta_2$, can be obtained graphically (if a series of volt-ampere characteristics of R_T is available), or analytically by substituting in (2) the value of R_a calculated from Eq. (4), and the $I_b(\theta)$ and $U_b'(\theta)$ values found from the general volt-ampere characteristic equation [2]

$$\frac{U}{I} = A \exp \frac{B}{\theta + UIF} , \tag{5}$$

where A, B, and F are the parameters of R_T.

With an accuracy sufficient for practical purposes, we can consider that $I_b(\theta)$ and $U_b(\theta)$ coincide with the coordinates of the volt-ampere characteristic maximum point; by using the condition for the existence of maxima $dU/dI = 0$, we obtain from (5)

$$I_b(\theta) = \sqrt{\frac{\alpha(\theta)}{F}}\left[\sqrt{A \exp \frac{B}{\theta + \alpha(\theta)}}\right]^{-1}; \quad U_b(\theta) = \sqrt{\frac{\alpha(\theta)}{F}}\sqrt{A \exp \frac{B}{\theta + \alpha(\theta)}}, \tag{6}$$

where

$$\alpha(\theta) = -\theta + \frac{B}{2} - \sqrt{\left(-\theta + \frac{B}{2}\right)^2 - \theta^2}.$$

For R_{T1}, Eqs. (5) and (6) are of the same form, but the A_1, B_1, and F_1 values are different.

Let us now formulate the conditions under which R_c can be kept constant throughout the entire $\theta_1 \le \theta \le \theta_2$ range. By substituting (6) in (3), it can be readily shown that R_c depends on θ (for a constant R_a) only due to a spread of the B values or, which is the same, due to a spread of the TR temperature coefficients, while R_c increases with a decrease in θ. Consequently, by assuming that $R_c = R_c(\theta_2)$ for $B_1 \ne B$, a temperature error is incurred, which is the larger, the higher the value of θ_2 in comparison with θ_1. The maximum allowable spread of B is found from the following condition [cf. (3)]:

$$R_C(\theta_2) = R_C^*(\theta_1 + \Delta\theta_1) = \frac{R_a U_{b1}(\theta_1 + \Delta\theta_1)}{U_b(\theta_1) - U_{b1}(\theta_1 + \Delta\theta_1)}, \tag{7}$$

where R_c^* is the value of R_c corresponding to the actuation temperature $\theta_1 + \Delta\theta_1$ for the setting temperature θ_1; $\Delta\theta_1 > 0$ is the maximum allowable error. After substituting in (7) the expression for $R_c(\theta_2)$ from (3) and the expressions for U_b and U_{b1} from (6), and after performing a number of calculations where the higher powers and products of the small quantities $\frac{B_1 - B}{B}$, $\frac{\theta}{B}$, and $\frac{\Delta\theta_1}{\theta_1}$ are neglected, we obtain the following value for the maximum allowable spread of B at the temperature θ_1:

$$\left[\frac{B_1 - B}{B}\right]_1 = \frac{\Delta\theta_1}{\theta_2 - \theta_1}. \tag{8}$$

If $R_c = R_c(\theta_1)$, the error will be of the opposite sign and, instead of (8), we obtain

$$\left[\frac{B_1 - B}{B}\right]_2 = \frac{-\Delta\theta_2}{\theta_2 - \theta_1}, \tag{9}$$

where $\Delta\theta_2 < 0$ is the maximum allowable error for the temperature setting θ_2, and $\left[\frac{B_1 - B}{B}\right]_2$ is the corresponding spread of B values.

It is obvious that the error will be smaller throughout the entire range if R_c remains within the range of $R_c = R_c(\theta_0)$ values, where $\theta_1 < \theta_0 < \theta_2$. Similarly to (8) and (9), we obtain, in this case,

$$\left[\frac{B_1 - B}{B}\right]_1 = \frac{\Delta\theta_1}{\theta_0 - \theta_1}, \quad \left[\frac{B_1 - B}{B}\right]_2 = \frac{-\Delta\theta_2}{\theta_2 - \theta_0}. \tag{10}$$

From symmetry requirements at the ends of the range, we have

$$\left[\frac{B_1 - B}{B}\right]_1 = \left[\frac{B_1 - B}{B}\right]_2 = \left[\frac{B_1 - B}{B}\right]_{12}, \quad \Delta\theta_1 = -\Delta\theta_2 = \Delta\theta_{12}. \tag{11}$$

From (10) and (11) we obtain

$$\theta_0 = \frac{\theta_1 + \theta_2}{2}, \quad \left[\frac{B_1 - B}{B}\right]_{12} = \frac{2\Delta\theta_{12}}{\theta_2 - \theta_1}. \tag{12}$$

The temperature coefficient spread of KMT-11 thermoresistors from a batch with the same rating usually does not exceed 0.3% per °C, which corresponds to a $\frac{B_1 - B}{B}$ spread of up to 0.06. If we assume that $\left[\frac{B_1 - B}{B}\right]_{12}$ = 0.06 for the range 35-100°C (θ_1 = 308°K, θ_2 = 373°K), we obtain from (12) $\theta_0 \sim 68$°C, $\Delta\theta_{12} \sim 2$°C. Thus, by using the R_c value for 68°C, the setting error can be reduced to ± 2°C throughout the entire 35-100°C range, without additional corrections for the spread of TR parameters; the error sign is determined during calibration by suitably shifting the setting points.

It should be noted that the setting error $\Delta\theta_{12}$ constitutes only a portion of the over-all error; its other portion consists of errors due to an inaccurate selection of R_a and R_c, an inaccurate adjustment of the feed voltage $U_f(\theta)$, changes in the load, etc. The latter errors also arise in the case of fixed setting and, under mass-production conditions, they can be reduced to ± 2-3°C; in this case, the over-all error will be 4-5°C. In certain cases, the indicated error may be unacceptable; then, in order to reduce $\Delta\theta_{12}$, it is necessary either to select a TR with respect to the temperature coefficient, or to average the TR characteristics by connecting in series several TR with a lower rating in a single thermal data transmitter.

8. On the basis of what has been said, the following conclusions can be drawn:

1) Thermal control devices, where the relay effect in circuits with TR is used, can be provided with adjustable actuation temperature setting (with an error of up to +5°C in the 35-100°C range) while preserving the miniaturization of the component units and the interchangeability of the relay units and thermal data transmitters.

2) The devices embodying these properties are flexible; they do not contain a great amount of wiring, and they are convenient in use and adjustment.

3) The construction of these devices does not entail the use of circuits and designs which would be greatly different from those presently available [1]; however, the requirements pertaining to the spread of TR parameters or, more accurately, the spread of TR temperature coefficients, must generally be more strict.

In conclusion, the author extends his thanks to G. K. Nechaev for the discussion of the paper, and his valuable advice.

LITERATURE CITED

1. G. K. Nechaev, Application of Thermoresistors in Temperature Control [in Russian] (Gostekhizdat, UkrSSR, Kiev, 1959).

2. G. N. Dul'nev, "Theory of thermal regimes for heat-sensitive semiconductor resistors," Collection: Devices and Regulators Used in Heat Engineering [in Russian] (Mashgiz, 1954).

EVENTS

SEMINAR-CONFERENCE ON THE THEORY AND METHODS OF MATHEMATICAL SIMULATION

G. M. Kozyreva

Translated from Avtomatika i Telemekhanika, Vol. 22, No. 1, pp. 125-126, January, 1961

The first session of the Seminar on the Theory and Methods of Mathematical Simulation, which was organized on the initiative of the Institute of Automation and Remote Control, Academy of Sciences of the USSR, was held on July 12-13, 1960 in Moscow. Problems pertaining to the design and application of digital simulators were discussed at the session.

More than 60 people from 17 scientific-research organizations in Moscow, Leningrad, and Gor'kii participated in the work of the Seminar. Five reports were heard and discussed.

In his introductory speech, B. Ya. Kogan noted that the Institute of Automation and Remote Control, Academy of Sciences of the USSR, considers that the solution of the question of organizing the discussion of the most important problems in computer techniques and, especially, in the field of their application in actual automatic systems, has been long overdue. In connection with this, the Institute intends to hold annual all-Union seminar-conferences. It is intended that each of these conferences be devoted to a certain important problem, and that it summarize the results achieved during the past year and indicate the direction of further development. B. Ya. Kogan further indicated that a situation has now arisen where the accuracy and the dynamic range of electronic simulating devices is insufficient for the solution of a number of important problems in the automatic control field. These limitations of electronic simulators are especially noticeable in solving complex (high-order) nonlinear differential equations, in the case where it is necessary to perform multiple coordinate transformations, and in the simulation of functions of a single and, especially, many variables.

The utilization of universal digital computers for these purposes is not justified because of the complexity of equipment and programming.

In recent years, a new method — digital simulation — has been developed, which apparently can be of great help in overcoming these difficulties.

A. V. Shileiko submitted a report entitled "A method for synthesizing the optimum structure of digital simulators." In his report, he proposed a method for synthesizing the optimum structure of specialized computers, which consists in the selection of the performance criterion, the determination of its dependence on a number of parameters, the so-called structural coefficients, which determine the structure of the device realizing an assigned algorithm, and in finding all the structural coefficient values securing the maximum performance criterion value. The amount of information provided by the device in unit time, reduced to the device complexity, is used as the performance criterion. This method was used for synthesizing the optimum structure of an actual digital simulator.

The report by L. M. Gol'denberg, entitled "Sequential digital differential analyzers (DDA)," was devoted to the description of the DDA "Integral" and provided some of its technical characteristics. The machine is of the sequential type and has a ferrite core memory. It has 32 integrators. The cyclic-pulse frequency is 50 kc. The iteration duration and, consequently, the integration speed, depend on the number of integrators fitted and, if their number is complete, it is 20 iterations per second. The machine contains 350 tubes and 8000 ferrite cores.

The report by K. S. Neslukhovskii, entitled "Integrating adapter to UTsVM," described a sequential-action DDA intended as a special adapter to the BÉSM-2 machine for solving differential equations. A standard magnetic drum memory is used in this device. The adapter operates at a rate of 50 iterations per second. In order to speed up the DDA operation, the programming of initial conditions is performed by means of a universal digital computer.

In his report, "Sequential DDA for higher frequencies," F. V. Maiorov reported the development of sequential DDA for operation in control systems.

In his communication, entitled "Frequency characteristics of DDA," S. V. Misailovskii proposed a method for determining the transmission band of DDA in dependence on their parameters, and he also determined the maximum actual rate of change in input signals produced without dynamic errors for a given frequency transmission band.

The following resolution was adopted at the seminar:

1. The initiative of IAT AN SSSR in organizing a permanently acting seminar on the theory and methods of mathematical simulation has been approved.

2. It is considered necessary to intensify efforts in designing digital simulators, to systematize the experience gained in their utilization, and to specify the fields of their possible application.

3. It is noted that problems pertaining to the methods and theory of digital simulation are not treated to a sufficient extent in the domestic technical literature.

Soviet Journals Available in Cover-to-Cover Translation

ABBREVIATION	RUSSIAN TITLE	TITLE OF TRANSLATION	PUBLISHER	TRANSLATION BEGAN Vol.	Issue	Year
AÉ	Atomnaya énergiya	Soviet Journal of Atomic Energy	Consultants Bureau	1	1	1956
Akust. zh.	Akusticheskii zhurnal	Soviet Physics – Acoustics	American Institute of Physics	1	1	1955
	Antibiotiki	Antibiotics	Consultants Bureau	4	1	1959
Astr(on). zh(urn).	Astronomicheskii zhurnal	Soviet Astronomy—AJ	American Institute of Physics	34	1	1957
Avto(mat). svarka	Avtomaticheskaya svarka	Automatic Welding	British Welding Research Association (London)		1	1959
	Avtomatika i Telemekhanika	Automation and Remote Control	Instrument Society of America	27	1	1956
	Biofizika	Biophysics	National Institutes of Health*		1	1957
	Biokhimiya	Biochemistry	Consultants Bureau	21	1	1956
Byull. éksp(erim). biol. i med.	Byulleten' éksperimental'noi biologii i meditsiny	Bulletin of Experimental Biology and Medicine	Consultants Bureau	41	1	1959
DAN (SSSR) Dokl(ady) AN SSSR	Doklady Akademii Nauk SSSR	The translation of this journal is published in sections, as follows:				
	Life Sciences	Doklady Biochemistry Section	American Institute of Biological Sciences	106	1	1956
		Doklady Biological Sciences Sections (Includes: Anatomy, biophysics, cytology, ecology, embryology, endocrinology, evolutionary morphology, genetics, histology, hydrobiology microbiology, morphology, parasitology, physiology, zoology sections)	American Institute of Biological Sciences	112	1	1957
		Doklady Botanical Sciences Sections (Includes: Botany, phytopathology, plant anatomy, plant ecology, plant embryology, plant physiology, plant morphology sections)	American Institute of Biological Sciences	112	1	1957
	Chemical Sciences	Proceedings of the Academy of Sciences of the USSR, Section: Chemical Technology	Consultants Bureau	106	1	1956
		Proceedings of the Academy of Sciences of the USSR, Section: Chemistry	Consultants Bureau	106	1	1956
		Proceedings of the Academy of Sciences of the USSR, Section: Physical Chemistry	Consultants Bureau	112	1	1957
	Earth Sciences	Doklady Earth Sciences Sections (Includes: Geochemistry, geology, geophysics, hydrogeology, mineralogy, paleontology, petrography, permafrost sections)	American Geological Institute	124	1	1959
		Proceedings of the Academy of Sciences of the USSR, Section: Geochemistry	Consultants Bureau	106-123	1 6	1957-1958
		Proceedings of the Academy of Sciences of the USSR, Sections: Geology	Consultants Bureau	106-123	1 6	1957-1958
	Mathematics	Doklady Soviet Mathematics	The American Mathematics Society	131	1	1961
	Physics	Soviet Physics—Doklady (Includes: Aerodynamics, astronomy, crystallography, cybernetics and control theory, electrical engineering, energetics, fluid mechanics, heat engineering, hydraulics, mathematical physics, mechanics, physics, technical physics, theory of elasticity sections)	American Institute of Physics	106	1	1956
		Proceedings of the Academy of Sciences of the USSR, Applied Physics Sections (does not include mathematical physics or physics sections)	Consultants Bureau	106-117	1	1956-1957
Derevoobrabat. prom-st'.	Derevoobrabatyvayushchaya promyshlennost'	Wood Processing Industry	Timber Development Association (London)		9	1959
	Élecktrosvyaz	Telecommunications	Massachusetts Institute of Technology*		1	1957
Éntom(ol). oboz(renie)	Éntomologicheskoe obozrenie	Entomological Review	American Institute of Biological Sciences	38	1	1959
Farmakol. (i) toksikol(ogiya)	Farmakologiya i toksikologiya	Pharmacology and Toxicology	Consultants Bureau	20	1	1957
FMM	Fizika metallov i metallovedenie	Physics of Metals and Metallography	Acta Metallurgica*	5	1	1957
Fiziol. zhurn. SSSR (im. Sechenova)	Fiziologicheskii zhurnal im. I. M. Sechenova	Sechenov Physiological Journal USSR	National Institutes of Health*		1	1957
Fiziol(ogiya) rast.	Fiziologiya rastenii	Plant Physiology	American Institute of Biological Sciences	4	1	1957
	Geokhimiya	Geochemistry	The Geochemical Society		1	1958
FTT	Fizika tverdogo tela	Soviet Physics—Solid State	American Institute of Physics	1	1	1959
Izmerit. tekh(nika)	Izmeritel'naya tekhnika	Measurement Techniques	Instrument Society of America		1	1959
Izv. AN SSSR, O(td). Kh(im). N(auk)	Izvestiya Akademii Nauk SSSR: Otdelenie khimicheskikh nauk	Bulletin of the Academy of Sciences of the USSR: Division of Chemical Sciences	Consultants Bureau		1	1952

continued

Izv. AN SSSR, O(td). T(ekhn). N(auk): Met(all). i top.	(see Met. i top.)					
Izv. AN SSSR Ser. fiz(ich).	Izvestiya Akademii Nauk SSSR: Seriya fizicheskaya	Bulletin of the Academy of Sciences of the USSR: Physical Series	Columbia Technical Translations		1	1954
Izv. AN SSSR Ser. geofiz.	Izvestiya Akademii Nauk SSSR: Seriya geofizicheskaya	Bulletin (Izvestiya) of the Academy of Sciences USSR: Geophysics Series	American Geophysical Union		1	1954
Izv. AN SSSR Ser. geol.	Izvestiya Akademii Nauk SSSR: Seriya geologicheskaya	Izvestiya of the Academy of Sciences of the USSR: Geologic Series	American Geological Institute		1	1958
Kauch. i rez.	Kauchuk i rezina	Soviet Rubber Technology	Research Association of British Rubber Manufacturers	18	1	1959
	Kinetika i kataliz	Kinetics and Catalysis	Consultants Bureau	1	3	1960
	Koks i khimiya	Coke and Chemistry USSR	Coal Tar Research Association (Leeds, England)		1	1958
Kolloidn. zh(urn).	Kolloidnyi zhurnal	Colloid Journal	Consultants Bureau	14	1	1952
	Kristallografiya	Soviet Physics – Crystallography	American Institute of Physics	2	1	1957
Metalov. i term. obrabot. metal.	Metallovedenie i termicheskaya obrabotka metallov	Metal Science and Heat Treatment of Metals	Acta Metallurgica	6	1	1958
	Metallurg	Metallurgist	Acta Metallurgica		1	1957
Met. i top.	Metallurgiya i topliva	Russian Metallurgy and Fuels	Eagle Technical Publications		1	1960
Mikrobiol.	Mikrobiologiya	Microbiology	American Institute of Biological Sciences	26	1	1957
OS	Optika i spektroskopiya	Optics and Spectroscopy	American Institute of Physics		6	1959
	Pochvovedenie	Soviet Soil Science	American Institute of Biological Sciences		1	1958
	Priborostroenie	Instrument Construction	British Scientific Instrument Research Association		1	1959
Pribory i tekhn. éks(perimenta)	Pribory i tekhnika éksperimenta	Instruments and Experimental Techniques	Instrument Society of America		1	1957
Prikl. matem. i mekh.	Prikladnaya matematika i mekhanika	Applied Mathematics and Mechanics	American Society of Mechanical Engineers		1	1958
PTÉ	(see Pribory i tekhn. éks.)					
	Problemy Severa	Problems of the North	National Research Council of Canada			
Radiotekh.	Radiotekhnika	Radio Engineering	Massachusetts Institute of Technology*	12	1	1957
Radiotekh. i élektronika	Radiotekhnika i élektronika	Radio Engineering and Electronics	Massachusetts Institute of Technology*	2	1	1957
	Stanki i instrument	Machines and Tooling	Production Engineering Research Assoc.		1	1959
	Stal'	Stal (in English)	Iron and Steel Institute		1	1959
Stek. i keram.	Steklo i keramika	Glass and Ceramics	Consultants Bureau	13	1	1956
Svaroch. proiz-vo	Svarochnoe proizvodstvo	Welding Production	British Welding Research Association		4	1959
Teor. veroyat. i prim.	Teoriya veroyatnostei i ee primenenie	Theory of Probability and Its Applications	Society for Industrial and Applied Mathematics		1	1956
Tsvet. Metally	Tsvetnye metally	Nonferrous Metals	Primary Sources		1	1960
UFN	Uspekhi fizicheskikh Nauk	Soviet Physics – Uspekhi (partial translation)	American Institute of Physics	66	1	1958
UKh	Uspekhi khimii	Russian Chemical Reviews	The Chemical Society (London)		1	1960
UMN	Uspekhi matematicheskikh nauk	Russian Mathematical Surveys	London Mathematical Society	15	1	1960
Usp. fiz. nauk	(see UFN)					
Usp. khim(ii)	(see UKh)					
Usp. matem. nauk	(see UMN)					
Usp. sovr. biol.	Uspekhi sovremennoi biologii	Russian Review of Biology	Oliver and Boyd		48	1959
Vest. mashinostroeniya	Vestnik mashinostroeniya	Russian Engineering Journal	Production Engineering Research Assoc.		4	1959
Vop. gem. i per. krovi	Voprosy gematologii i perelivaniya krovi	Problems of Hematology and Blood Transfusion	National Institutes of Health*		1	1957
Vop. onk.	Voprosy onkologii	Problems of Oncology	National Institutes of Health*		1	1957
Vop. virusol.	Voprosy virusologii	Problems of Virology	National Institutes of Health*		1	1957
Zav(odsk). lab(oratoriya)	Zavodskaya laboratoriya	Industrial Laboratory	Instrument Society of America	25	1	1959
ZhAKh Zh. anal(it). khimii	Zhurnal analiticheskoi khimii	Journal of Analytical Chemistry USSR	Consultants Bureau	7	1	1952
ZhETF Zh. éksperim. i teor. fiz.	Zhurnal éksperimental'noi i theoreticheskoi fiziki	Soviet Physics–JETP	American Institute of Physics	28	1	1955
ZhFKh Zh. fiz. khimii	Zhurnal fizicheskoi khimii	Russian Journal of Physical Chemistry	The Chemical Society (London)		7	1959
ZhMÉI Zh(urn). mikrobiol. épidemiol. i immunobiol.	Zhurnal mikrobiologii, épidemiologii i immunobiologii	Journal of Microbiology, Epidemiology and Immunobiology	National Institutes of Health*		1	1957
ZhNKh Zh(urn). neorgan(ich). khim(ii)	Zhurnal neorganicheskoi khimii	The Russian Journal of Inorganic Chemistry	The Chemical Society (London)		1	1959
ZhOKh Zh(urn). obshch(ei) khimii	Zhurnal obshchei khimii	Journal of General Chemistry USSR	Consultants Bureau	19	1	1949
ZhPKh Zh(urn). prikl. khimii	Zhurnal prikladnoi khimii	Journal of Applied Chemistry USSR	Consultants Bureau	23	1	1950
ZhSKh Zh(urn). strukt. khimii	Zhurnal strukturnoi khimii	Journal of Structural Chemistry	Consultants Bureau	1	1	1960
ZhTF Zh(urn). tekhn. fiz.	Zhurnal teknicheskoi fiziki	Soviet Physics–Technical Physics	American Institute of Physics	26	1	1956
Zh(urn). vyssh. nervn. deyat. (im. Pavlova)	Zhurnal vysshei nervnoi deyatel'nosti (im. I. P. Pavlova)	Pavlov Journal of Higher Nervous Activity	National Institutes of Health*		1	1958

*Sponsoring organization. Translation through 1960 issues is a publication of Pergamon Press.